

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 14:33:42 ; Search time 3672.39 Seconds

(without alignments)  
11294.891 Million cell updates/sec

Title: US-10-076-747-7

Perfect score: 957

Sequence: 1 ttggagcttaccatcatcatg.....gagcagagagccagagct 957

Scoring table: IDENTITY NJC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

1: gb\_ba.\*

2: gb\_hcg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_pl.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_hcg\_hum.\*

31: em\_hcg\_inv.\*

32: em\_hcg\_other.\*

33: em\_hcg\_mus.\*

34: em\_hcg\_pln.\*

35: em\_hcg\_rtd.\*

36: em\_hcg\_mam.\*

37: em\_hcg\_vrt.\*

38: em\_sy.\*

39: em\_hcgo\_hum.\*

40: em\_hcgo\_mus.\*

41: em\_hcgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56.4	5.9	991	11	PM12H12B
2	53.8	5.6	7238	6	1664455 Penicillium
3	53.4	5.6	218201	2	AC141731 Apis mell
4	53	5.5	1392	11	PM12B10G
5	53	5.5	160634	2	AC141999 Rattus no
6	53	5.5	177485	2	AC090975 Mus muscu
7	50.8	5.3	10732	6	E32986 Gene encodi
8	50.8	5.3	125020	2	AF429315 Homo sapi
9	50.4	5.3	123570	2	AC142034 Rattus no
10	50.4	5.3	125020	9	AF429315 Homo sapi
11	50.2	5.2	46251	9	AD000671 Homo sapi
12	50.2	5.2	194835	2	AC022315 Homo sapi
13	50	5.2	152666	2	AC143344 Macaca mu
14	49.6	5.2	2000	6	AX655393 Sequence
15	49.6	5.2	142852	8	AF161259 Oryza sat
16	49.6	5.2	163032	2	AC144742 Oryza sat
17	49.6	5.2	187223	2	AC008333 Drosophila
18	49.4	5.2	1393	11	PM1H12G
19	49.2	5.1	189230	2	AC120836 Mus muscu
20	49	5.1	177877	2	AC13243 Rattus no
21	48.8	5.1	85434	2	AC066610 Homo sapi
22	48.4	5.1	1070	9	ORANGIT
23	48.2	5.0	166352	9	AC068530 Homo sapi
24	48.2	5.0	177055	9	AC090360 Homo sapi
25	48.2	5.0	276546	9	AC139100 Homo sapi
26	47.6	5.0	150942	2	AC130463 Homo sapi
27	47.4	5.0	576	9	HS335249 Homo sapi
28	47	4.9	72847	9	AC023563 Homo sapi
29	47	4.9	197082	9	AC006249 Homo sapi
30	47	4.9	239130	2	AC079420 Mus muscu
31	47	4.9	256673	2	AC087146 Homo sapi
32	46.8	4.9	140176	2	AC016411 Homo sapi
33	46.8	4.9	160587	8	OSUN00222 Oryza sat
34	46.6	4.9	39954	9	AP006257 Homo sapi
35	46.4	4.8	125402	9	AC114815 Homo sapi
36	46.4	4.8	136551	2	AC048354 Homo sapi
37	46.4	4.8	202058	2	AC016913 Homo sapi
38	46	4.8	1065	11	PM2B12B
39	46	4.8	63082	2	AC022663 Homo sapi
40	46	4.8	64189	2	AC015511 Homo sapi
41	46	4.8	110000	2	AC143301 Macaca mu
42	46	4.8	156289	2	AC144155 Macaca mu
43	46	4.8	262074	2	AC098624 Rattus no
44	46	4.8	272166	2	AC120907 Rattus no
45	45.8	4.8	83569	2	AC022967 Homo sapi

## ALIGNMENTS

RESULT 1

PM12H12B

LOCUS Penicillium marneffei STS, clone pm12h12.b, sequence tagged site.

DEFINITION

ACCESSION AF684455

VERSION

KEYWORDS

SOURCE

ORGANISM

STX

Penicillium marneffei

Penicillium marneffei

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

Eurotiaceae; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.

REFERENCE

1 Yuen,K.Y., Pascal,G., Wong,S.S., Glaeser,P., Woo,P.C., Kunst,F.,

Cal,J.O., Cheung,E.Y., Medigue,C. and Danchin,A.

Exploring the Penicillium marneffei genome

JOURNAL Arch. Microbiol. 179 (5), 339-353 (2003)  
 MEDLINE 22595073  
 PUBMED 12640520  
 REFERENCE 2 (bases 1 to 991)  
 AUTHORS Danchin, A. and Pascal, G.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-MAR-2002) Danchin A., HKU-Pasteur Research Centre,  
 Dexter HC Man Building 8, Sarsason Road, Pokfulam, Hong Kong

FEATURES  
 Source  
 1. .991  
 /organism="Penicillium marneffei"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:37727"  
 /clone="pm12h12.b"

ORIGIN  
 Query Match 5.2%; Score 56.4; DB 11; Length 991;  
 Best Local Similarity 50.2%; Pred. No. 0.039;  
 Matches 135; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

685 CCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 744  
 Db CCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 539  
 480 CCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 539  
 745 ACAATACGAGAGCGGAGGACCGGCGGCGGAGAGCCCACTGGGCGGCGGAGGAG 804  
 Db GCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 599  
 540 GCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 599  
 805 GCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 864  
 Db GCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 659  
 600 GCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 659  
 865 GCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 924  
 Db GCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 719  
 660 GCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 719  
 925 CCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 953  
 Db CCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 748  
 720 CCGGAGACCGGCGGAGGACCCCGGCGGAGAGCCCACTGGGCGGCGGAGGAGG 748

RESULT 2  
 166494/c  
 LOCUS 166494 7218 bp DNA linear PAT 28-DEC-1997  
 DEFINITION Sequence 14 from patent US 5670367.  
 ACCESSION 166494  
 VERSION 166494.1 GI:2724471  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 7218)  
 AUTHORS Dornier, F., Scheiflinger, F. and Falkner, F. Gunter.  
 TITLE Recombinant fowlpox virus  
 JOURNAL Patent: US 5670367-A 14 23-SEP-1997;  
 FEATURES  
 Location/Qualifiers  
 1. .7218  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Query Match 5.6%; Score 53.8; DB 6; Length 7218;  
 Best Local Similarity 1.9%; Pred. No. 0.13;  
 Matches 7; Conservative 219; Mismatches 141; Indels 0; Gaps 0;

581 ATGGCGTACCGGCGGATCAATTGCGCGGAGCTACAGACTTAAAGTTACCGCGCA 640  
 Db ATGGCGTACCGGCGGATCAATTGCGCGGAGCTACAGACTTAAAGTTACCGCGCA 1384  
 1443 ATTTGGTACCGGCGGATCAATTGCGCGGAGCTACAGACTTAAAGTTACCGCGCA 1384  
 641 GCGGCGGAGAGAGCGGCTTCCGAGGAGTCCCAACAGAGGCGCGGAGCGGAGCGC 700  
 Db GCGGCGGAGAGAGCGGCTTCCGAGGAGTCCCAACAGAGGCGCGGAGCGGAGCGC 1324  
 1383 RR 1324

QY 701 GACCCCGGCGGAGAGACCCCACTGCGGCGGAGGCGGAGGAGAGGAGGAGGAGG 760  
 Db RR 1264  
 1323 RR 1264  
 QY 761 GACCCAGACCGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 820  
 Db RR 1204  
 1263 RR 1204  
 QY 821 GCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 880  
 Db RR 1144  
 1203 RR 1144  
 QY 881 GCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 940  
 Db RR 1084  
 1143 RR 1084  
 QY 941 AGCAGAG 947  
 Db RR 1084  
 1083 RR 1084

RESULT 3  
 AC141731  
 LOCUS AC141731  
 DEFINITION Apis mellifera clone CH224-58F22, WORKING DRAFT SEQUENCE, 91  
 ACCESSION AC141731  
 VERSION AC141731.1 GI:29123915  
 KEYWORDS HTGS, HTGS PHASE1, HTGS DRAFT.  
 SOURCE Apis mellifera (honeybee)  
 ORGANISM Apis mellifera  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
 Apoidea; Apis.  
 1 (bases 1 to 218201)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., All-ouman, F.R., Allen, C.,  
 Alsbrooks, S.L., Amaral, H.C., Are, O.R., Ayale, M., Banks, T.,  
 Barbarella, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,  
 Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burck, P., Burkett, C., Byrd, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
 Homsl, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Koryah, J., Kovar, C.,  
 Kratovic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louleghed, H.,  
 Lora, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawlin, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabadi, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokwenkwo, S., Oguh, M., Okunolu, G.,  
 Oreguine, N., Oyedele, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primm, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rivers, M., Rojas, A., Rojurokan, I., Rolfe, M., Ruiz, S., Severy, G.,  
 Scherer, S., Scott, G., Shen, H., Shoshitaishvili, S., Sisson, I.,  
 Sodergren, E., Sonakke, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tameis, A., Tameis, A., Tang, R.,  
 Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Umani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S.,  
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

## COMMENT

Mu C., Mu Y., Mu Y.F., Zhou J., Zorrilla S., Nelson D.,  
Weinstock G. and Gibbs R.  
Direct Submission  
Unpublished  
2 (bases 1 to 218201)  
Worley K.C.  
Submitted  
Submitted (19-MAR-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: AMEV  
Center clone name: CH224-58F22  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye 1.00% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 195857 bases at least Q40  
Consensus quality: 207269 bases at least Q30  
Consensus quality: 213633 bases at least Q20  
Estimated insert size: 183061; sum-of-contigs estimation  
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 91 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
1459: contig of 1459 bp in length  
1460 1559: gap of unknown length  
1560 2927: contig of 1368 bp in length  
2928 3027: gap of unknown length  
3028 4064: contig of 1037 bp in length  
4065 4164: gap of unknown length  
4165 5450: contig of 1286 bp in length  
5451 5550: gap of unknown length  
5551 6910: contig of 1360 bp in length  
6911 7010: gap of unknown length  
7011 8455: contig of 1445 bp in length  
8456 8555: gap of unknown length  
8556 9628: contig of 1073 bp in length  
9629 9728: gap of unknown length  
9729 10754: contig of 1026 bp in length  
10755 10854: gap of unknown length  
10855 12039: contig of 1185 bp in length  
12040 12139: gap of unknown length  
12140 13142: contig of 1003 bp in length  
13143 13242: gap of unknown length  
13243 14333: contig of 1091 bp in length  
14334 14433: gap of unknown length  
14433 15532: contig of 1099 bp in length  
15532 15632: gap of unknown length  
15633 16838: contig of 1206 bp in length  
16839 16938: gap of unknown length  
16939 18146: contig of 1208 bp in length  
18147 18246: gap of unknown length  
18247 19466: contig of 1220 bp in length  
19467 19566: gap of unknown length  
19567 20905: contig of 1339 bp in length  
20906 21005: gap of unknown length  
21006 22525: contig of 1520 bp in length  
22526 23739: contig of 1114 bp in length  
23739: contig of 1114 bp in length

23740 23840 23849: gap of unknown length  
23840 24845: contig of 1006 bp in length  
24846 24945: gap of unknown length  
24946 26300: contig of 1355 bp in length  
26301 26400: gap of unknown length  
26401 27643: contig of 1243 bp in length  
27644 27744: gap of unknown length  
27744 28940: contig of 1197 bp in length  
28941 29040: gap of unknown length  
29041 30245: contig of 1205 bp in length  
30246 30345: gap of unknown length  
30346 31518: contig of 1173 bp in length  
31519 31618: gap of unknown length  
31619 33173: contig of 1555 bp in length  
33174 33273: gap of unknown length  
33274 34738: contig of 1465 bp in length  
34739 34838: gap of unknown length  
34839 35982: contig of 1144 bp in length  
35983 36082: gap of unknown length  
36083 37444: contig of 1362 bp in length  
37445 37544: gap of unknown length  
37545 38752: contig of 1208 bp in length  
38753 38852: gap of unknown length  
38853 40154: contig of 1302 bp in length  
40155 40254: gap of unknown length  
40255 42036: contig of 1782 bp in length  
42037 42137: gap of unknown length  
42138 43361: contig of 1225 bp in length  
43362 43461: gap of unknown length  
43462 44861: contig of 1400 bp in length  
44862 44961: gap of unknown length  
44962 46359: contig of 1388 bp in length  
46360 46459: gap of unknown length  
46460 47564: contig of 1105 bp in length  
47565 47664: gap of unknown length  
47665 49513: contig of 1849 bp in length  
49514 49613: gap of unknown length  
49614 50721: contig of 1108 bp in length  
50722 50821: gap of unknown length  
50822 52553: contig of 1732 bp in length  
52554 52653: gap of unknown length  
52654 54428: contig of 1775 bp in length  
54429 54529: gap of unknown length  
54530 55799: contig of 1271 bp in length  
55799 55899: gap of unknown length  
55900 56993: contig of 1094 bp in length  
56994 57093: gap of unknown length  
57094 58796: contig of 1603 bp in length  
58797 60055: contig of 1259 bp in length  
60056 60155: gap of unknown length  
60156 61586: contig of 1431 bp in length  
61587 63051: contig of 1365 bp in length  
63052 63151: gap of unknown length  
63152 64599: contig of 1448 bp in length  
64600 64699: gap of unknown length  
64700 66345: contig of 1846 bp in length  
66346 66645: gap of unknown length  
66646 68608: contig of 1963 bp in length  
68609 68708: gap of unknown length  
68709 70013: contig of 1305 bp in length  
70014 70113: gap of unknown length  
70114 71424: contig of 1311 bp in length  
71425 71524: gap of unknown length  
71525 73106: contig of 1582 bp in length  
73107 73206: gap of unknown length  
73207 74473: contig of 1267 bp in length  
74474 74573: gap of unknown length  
74574 76648: contig of 2075 bp in length  
76649 76749: gap of unknown length  
76750 78054: contig of 1306 bp in length  
78055 78154: gap of unknown length





TITLE  
 JOURNAL  
 REFERENCES  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Zhao S., Dunn D., von Niederhausern A., Weiss R., Smith D.R.,  
 Holt R.A., Smith H.O., Weinstock G. and Gibbs R.A.  
 Direct Submission  
 2 (bases 1 to 160634)  
 Morley K.C.  
 Submitted (21-MAR-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GRBY  
 Center clone name: CH230-270C13  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 126970 bases at least Q40  
 Consensus quality: 134664 bases at least Q30  
 Consensus quality: 140158 bases at least Q20  
 Estimated insert size: 126396; sum-of-coverage estimation  
 Quality coverage: 2x in Q20 bases; sum-of-coverage estimation

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Gembank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Gembank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 61 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1002: contig of 1002 bp in length  
 \* 1003 1102: gap of unknown length  
 \* 1103 2104: contig of 1002 bp in length  
 \* 2105 2204: gap of unknown length  
 \* 2205 3741: contig of 1537 bp in length  
 \* 3742 3841: gap of unknown length  
 \* 3842 5455: contig of 1614 bp in length  
 \* 5456 5555: gap of unknown length  
 \* 5556 7319: contig of 1764 bp in length  
 \* 7320 7419: gap of unknown length  
 \* 7420 8494: contig of 1075 bp in length  
 \* 8495 8594: gap of unknown length  
 \* 8595 10312: contig of 1718 bp in length  
 \* 10313 10412: gap of unknown length  
 \* 10413 11608: contig of 1196 bp in length  
 \* 11609 11708: gap of unknown length  
 \* 11709 12796: contig of 1088 bp in length  
 \* 12797 12896: gap of unknown length  
 \* 12897 14011: contig of 1115 bp in length  
 \* 14012 14111: gap of unknown length  
 \* 14112 15372: contig of 1261 bp in length  
 \* 15373 15472: gap of unknown length  
 \* 15473 16702: contig of 1230 bp in length  
 \* 16703 16802: gap of unknown length  
 \* 16803 18540: contig of 1738 bp in length  
 \* 18541 18640: gap of unknown length  
 \* 18641 19848: contig of 1208 bp in length  
 \* 19849 19948: gap of unknown length  
 \* 19949 21085: contig of 1137 bp in length  
 \* 21086 21185: gap of unknown length  
 \* 21186 22538: contig of 1353 bp in length  
 \* 22539 22639: gap of unknown length  
 \* 22639 24199: contig of 1561 bp in length  
 \* 24200 24299: gap of unknown length  
 \* 24300 25490: contig of 1191 bp in length

25491 25590: gap of unknown length  
 \* 25591 26810: contig of 1220 bp in length  
 \* 26811 26910: gap of unknown length  
 \* 26911 28333: contig of 1423 bp in length  
 \* 28334 28433: gap of unknown length  
 \* 28434 29824: contig of 1391 bp in length  
 \* 29825 29924: gap of unknown length  
 \* 29925 31363: contig of 1439 bp in length  
 \* 31364 31463: gap of unknown length  
 \* 31464 33704: contig of 2241 bp in length  
 \* 33705 33804: gap of unknown length  
 \* 33805 35873: contig of 2069 bp in length  
 \* 35874 35973: gap of unknown length  
 \* 35974 38042: contig of 2069 bp in length  
 \* 38043 38142: gap of unknown length  
 \* 38143 39581: contig of 1439 bp in length  
 \* 39582 39681: gap of unknown length  
 \* 39682 41011: contig of 1330 bp in length  
 \* 41012 41111: gap of unknown length  
 \* 41112 42348: contig of 1237 bp in length  
 \* 42349 42448: gap of unknown length  
 \* 42449 44108: contig of 1660 bp in length  
 \* 44109 44208: gap of unknown length  
 \* 44209 47160: contig of 2352 bp in length  
 \* 47161 47260: gap of unknown length  
 \* 47261 49178: contig of 1918 bp in length  
 \* 49179 49278: gap of unknown length  
 \* 49279 51273: contig of 1995 bp in length  
 \* 51274 51373: gap of unknown length  
 \* 51374 53195: contig of 1822 bp in length  
 \* 53196 53295: gap of unknown length  
 \* 53296 55875: contig of 2580 bp in length  
 \* 55876 55975: gap of unknown length  
 \* 55976 58197: contig of 2222 bp in length  
 \* 58198 58297: gap of unknown length  
 \* 58298 60391: contig of 2094 bp in length  
 \* 60392 60491: gap of unknown length  
 \* 60492 61807: contig of 1316 bp in length  
 \* 61808 61907: gap of unknown length  
 \* 61909 64495: contig of 2588 bp in length  
 \* 64496 64595: gap of unknown length  
 \* 64596 67408: contig of 2813 bp in length  
 \* 67409 67508: gap of unknown length  
 \* 67509 69971: contig of 2463 bp in length  
 \* 69972 70071: gap of unknown length  
 \* 70072 72989: contig of 2918 bp in length  
 \* 72990 73089: gap of unknown length  
 \* 73090 74939: contig of 1850 bp in length  
 \* 74940 75039: gap of unknown length  
 \* 75040 76951: contig of 1912 bp in length  
 \* 76952 77051: gap of unknown length  
 \* 77052 80255: contig of 3204 bp in length  
 \* 80256 80355: gap of unknown length  
 \* 80356 82038: contig of 1683 bp in length  
 \* 82039 82138: gap of unknown length  
 \* 82139 85079: contig of 2941 bp in length  
 \* 85080 85179: gap of unknown length  
 \* 85180 88512: contig of 3333 bp in length  
 \* 88513 88612: gap of unknown length  
 \* 88613 92252: contig of 3640 bp in length  
 \* 92253 92352: gap of unknown length  
 \* 92353 94705: contig of 2353 bp in length  
 \* 94706 94805: gap of unknown length  
 \* 94806 98359: contig of 3554 bp in length  
 \* 98360 98459: gap of unknown length  
 \* 98460 103311: contig of 4852 bp in length  
 \* 103312 103411: gap of unknown length  
 \* 103412 107829: contig of 4418 bp in length  
 \* 107830 107929: gap of unknown length

Query Match 5.5%; Score 53; DB 2; Length 160634;  
 Best Local Similarity 53.0%; Pred. No. 0.16;  
 Matches 133; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

RESULT 6			
AC090975/c			
LOCUS	AC090975	177485 bp	DNA linear HTG 15-MAY-2001
DEFINITION	Mus musculus chromosome 17 clone RP23-290119 strain C57BL6/J,		
	WORKING DRAFT SEQUENCE, 15 unordered pieces.		
ACCESSION	AC090975		
VERSION	AC090975.8	GI:1955112	
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1 (bases 1 to 177485)		
	Gillis,G., Li,L., Montgomery,R.T., Chin,D., Decker,J., Fusina,M.,		
	Goltz,J., Haider,A., Hall,I., Han,J., Ioshikhes,I.P., Lee,E.,		
	Perezera,A., Shim,C., Thomas,E. and Kuchelapatti,R.		
TITLE	BperA,Throughput Mouse Sequencing		

TITLE	High Throughput Mouse Sequencing
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 177485)
AUTHORS	Grill-B., Li,L., Montgomery,K.T., Chiu,D., Decker,J., Fusina,M., Goltz,J., Halder,A., Hall,L., Han,J., Ioshikhes,I.P., Lee,E., Pereira,A., Shim,C., Thomas,E. and Kucherlapati,R.
TITLE	Direct Submission
JOURNAL	Submitted (22-MAR-2001) Department of Molecular Genetics, Albert Einstein College of Medicine Genome Center, 1300 Morris Park Ave. Bronx, NY 10461, USA
REFERENCE	3 (bases 1 to 177485)
AUTHORS	Grill-B., G., Li,L., Montgomery,K.T., Chiu,D., Decker,J., Fusina,M., Goltz,J., Halder,A., Hall,L., Han,J., Ioshikhes,I.P., Lee,E., Pereira,A., Shim,C., Thomas,E. and Kucherlapati,R.
TITLE	Direct Submission
JOURNAL	Submitted (20-MAR-2002) Harvard Partners Center for Genetics and Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA 02139, USA
COMMENT	On Mar 20, 2002 this sequence version replaced gi:17486574. -----Genome Center Center: Harvard Partners Genome Center Web site: http://www.hpcgs.org/Sequence/mouse.html Contact: hpgc@mednet.mgh.harvard.edu -----Summary Statistics Center project name: AEU Sequencing vector: pUC18; L08752 Chemistry: Dye-terminator Big Dye; 100% *Consensus quality: 17252 at least Q20 *Consensus quality: 171774 at least Q30 *Consensus quality: 170494 at least Q40 Estimated insert size: agarose-FP - N/A **estimated insert size: 17705 - sum-of-contigs Quality coverage: 10.8 x in Q20 bases, sum-of-contigs estimation

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
*
* 1
* 31079: contig of 31079 bp in length
*
* 31080 31099: gap of unknown length
* 68493: contig of 37394 bp in length
* 31100 68494: gap of unknown length
* 68494 68513: gap of unknown length
* 68514 99930: contig of 31417 bp in length
* 99931 99950: gap of unknown length
* 99951 130275: contig of 30325 bp in length
* 130276 130295: gap of unknown length
* 130296 149687: contig of 19392 bp in length
* 149688 149707: gap of unknown length
* 149708 159762: contig of 10055 bp in length
* 159763 159783: gap of unknown length
* 159784 167610: contig of 7828 bp in length
* 167611 167630: gap of unknown length
* 167631 170306: contig of 5406 bp in length
* 170307 173056: gap of unknown length
* 173057 173359: contig of 303 bp in length
* 173360 173379: gap of unknown length
* 173380 174675: contig of 1296 bp in length
* 174676 174695: gap of unknown length
* 174696 175576: contig of 881 bp in length
* 175577 175596: gap of unknown length
* 175597 175722: contig of 126 bp in length
* 175723 175742: gap of unknown length
* 175743 176698: contig of 956 bp in length
* 176699 176718: gap of unknown length
* 176719 177383: contig of 665 bp in length
* 177384 177403: gap of unknown length
* 177404 177485: contig of 82 bp in length.
*
* Location/Qualifiers
*
* source
*
* 1. 177485
*
*   /organism="Mus musculus"
*   /mol_type="genomic DNA"
*   /strain="C57BL6/J"
*   /db_xref="taxon:10090"
*   /chromosome="17"
*   /clone="RP23-290119"
*   /sex="male"
*
* 1. 31079
*
*   /note="assembly_name:Contig161"
*   clone_end:576
*   vector_side:right"
*   31100..68493
*   /note="assembly_name:Contig160"
*   clone_end:r7
*   vector_side:left"
*
* 68514..99930
*   /note="assembly_name:Contig159"
*   99951..130275
*   /note="assembly_name:Contig158"
*   130286..149687
*   /note="assembly_name:Contig157"
*   149708..159762
*   /note="assembly_name:Contig156"
*   159783..167610
*   /note="assembly_name:Contig155"
*   167631..173056
*   /note="assembly_name:Contig154"
*   173057..173359
*   /note="assembly_name:Contig153"
*   173380..174675
*   /note="assembly_name:Contig152"
*   174696..175576
*   /note="assembly_name:Contig151"

```

```
misc_feature      175572      175587      .assembly_name:Contig150"
misc_feature      175743      176698      /note="assembly_name:Contig149"
misc_feature      176719      177353      /note="assembly_name:Contig148"
misc_feature      177404      177465      /note="assembly_name:Contig147"
ORIGIN
```

Query Match	Best Local Similarity	5.5%;	Score 53;	DB 2;	Length 177485;
Best Local Similarity	49.8%;	Pred. No. 0.16;			
Matches 164;	Conservative 0;	Mismatches 160;	Indels 5;	Gaps 1;	
QY	628	AGTTGACCGCGCAGCGGCCGGAAGACAGGCTTTTCAGCGGATATCCAAACACAGAGGCCG	687		
Db	68247	AGCCGCGCCCGCAGCGGAGACGCGCGCCGCGCGGGGGGGCGAGACCGACCGCGGGCGG	68188		
QY	688	CAGACCGGCGACGCGACCCCGGGGGGAGAGACCCCACTGCGCGGACGAGCGCGAGCGGAAAGACA	747		
Db	68187	CGCGGGCGCCCGCGCACGCGGAGAGGGGGGGGGCGCGCGCCACGCGGGGGGACACCG	68128		
QY	748	GATACGAGGACGCGGAGCCACACGCGCGCCCGTCATAGACGAGACGG-----CGAGG	802		
Db	68127	CGGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGGAGGACGAGGAGGGGCGGGACCGCGAGG	6806		
QY	803	GAGCGAGACCCCGAAGCAGCGCGCCCAAGACGACCGCGCAAGCCACACGCGCCTAGAGCC	862		
Db	68067	GCGGCGCCCGCGCGGAGCGCGCGGGGGGGGGCGGGGGGCGGCCACCCCGCGCG	68008		
QY	863	TGCGCCACGCGCGCGCCACGCGCGCGAGAGCGGGGCGGAGCACCGCAGAGACCGACCA	922		
Db	68007	CGCGCGCCCGCGCGCGCCCGCCCGCACCGCGCCGCAACCCACCGCCACACCCACGACCC	67948		
QY	923	ACCAGCGGCGCCACGCGGAGAGGCGAC	951		
Db	67947	GCCCCCGGACGACGCGCGCCCGACGCCAC	67919		

RESULT	7
E32986	
LOCUS	E32986 10732 bp DNA linear PAT 18-JUN-2001
DEFINITION	Gene encoding cellulose synthetaser.
ACCESSION	E32986
VERSION	E32986.1 GI:13022340
KEYWORDS	JP 2000060568-A/1..
SOURCE	Vigna angularis (adzuki bean)
ORGANISM	Vigna angularis Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna
REFERENCE	1 (bases 1 to 10732) Koichi,M.,Tomoniko,K., Shigenru,S. and Daisuke,S. Title Gene encoding cellulose synthetaser Journal Patent: JP 2000060568-A 1 29-FEB-2000, KOICHI MIZUNO,MITSUJI GUYOUSAI SHOKUETSU BIO KENKUSHO OS Vigna angularis
COMMENT	

```

PF      26-AUG-1998 JP 1598235998
PR
PI      KOIHI MIZUNO, TOMOHIKO KATO, SHIGERU SATO, DAISUKE SHIBATA PC
C12N15/09,A01H5/00,C07K16/40,C12N5/10,C12N9/10,C12P21/02// PC
(C12N15/09,C12R1:91), (C12N5/10,C12R1:91), (C12P21/02,C12R1:91), PC
C12N15/00,
PC      C12N5/00, (C12N15/00,C12R1:91), (C12N5/00,C12R1:91) CC
FH      Key      Location/Qualifiers
FT      CDS      Location/Qualifiers
              (1)..(3375) .

FEATURES
      source
      1..10732
      /organism="Vigna angularis"
      /mol_type="genomic DNA"

```

```

ORIGIN                                     /db_xref="taxon:3914"
Query Match                               5.3%; Score 50.8; DB 6; Length 10732;
Best Local Similarity 14.8%; Pred. No. 0.59;
Matches 135; Conservative 333; Mismatches 445; Indels 2; Gaps 2;

```

Qy	41	GATGATGCGTGGTCCGGCCGAGGATGACGGGACACTCTTTCGAGCTCAAGGCTG	100
Db	9578	GASHHARGRGRTSYSGRSKRYTSRSVAGTAASRBSASMTASASVAAWGTYSAAV	9637
Qy	101	ATGCAAGTGTGAGTGATCCACAGAACTACAGCCTGCAGAGCTTTAGTGCATGAAGG	160
Db	9638	AVASRVAAGSYAAATHRGNSRRSPAASTRYSAAVATHRASRSRGRTHRGTHYSGRTRR	9697
Qy	161	AGACACCGGTGATGAAGACAGTGGTCCCTACAGAAATGTTCTATAGGTTCTTAACAG	220
Db	9698	RGGGYTSGRVAAASHSRGNTHTHRTGTRTHRGVAAVAAWMTGNTSASHSRBSRY	9757
Qy	221	CTCAGCCCCCACTACCAATGGCAGACCTAGCACGCTGACGAGGATCCCAAGAGAGGGT	280
Db	9758	SSRYSASNTHRCYSRASHYSGTTHRVAGYRGNAVSVAASTRGNAASHVATHH	9817
Qy	281	CTCTCCATCCACACCAAGGCGAGTCAAAGCCCTTATCATCGCATGCTCCAGCTC	340
Db	9818	SYSGYRASASHSHSYVAALAAKSHVAAASHASMAHSAASHSRARRGYSRAVASHH	9877
Qy	341	ATGTAAAGCGCCTTACAAATGAATGATTTCTGCACTTGTTGAATGTTCTTACATATTA	400
Db	9878	RGNASHRRGNCTSHGCTSYSSRCYSRBSRSRVAAVAAASHNRSPAAGXASVAAASRYSAG	9937
Qy	401	CAAAAGACACTACTCACTACACGAGTGCATGCATCCGGTCCGTGGCGGCAATGCC	460
Db	9938	GRSRGNAGS-SRYSRSGYSGVAGAGRAAVAASTYRASRSRBSAGNTSGYVAGASH	9996
Qy	461	ACTTCGCTCCGCGCTCCATCATATATACATATGTTACACACCGGTGTCCGCGCACAG	520
Db	9997	SRTRYASASGMYTTRTHRGTYRTMTVAAGVAAHGNTRYTHHSRGRYSVAAHSHVAC	10056
Qy	521	TATGATGATATTTCCGCCCGGATCCAGACCATGGATTTCCCCCGTGGCTGATCGAAT	580
Db	10057	STYRASRGRYSTHRASTYRGVASHNAAYTHTHTHAASTMTBASNHSCYHSRGNAY	10116
Qy	581	ATCGGGAACCGGCATCAATTCGCGCGGAGAGTACAGACTTAAATAATGACCGCGCA	640
Db	10117	ASHSRMTARASAAIYSGNHASTRASTGTHPARGSASVGGASNTGHHSHVAGISGANG	10176
Qy	641	GCGGCGGAAGACAGGCTTTGACGCGAATGCCAAACAGAGGCCGACGACCGCGAGCC	700
Db	10177	HSBARGASNSRRRGNVATHRTHSRCYSGHSAAGNSAGNSAGNSAGNYGTYRTYRANNGYSGHTH	10236
Qy	701	GACCGCGGAGGAGAGAGCCCATCTGGGAGGCGGAGGAAGAAAGATGACAGAGCC	760
Db	10237	RATYSASYSRRAAAGSRGNITSVAVGAGTSGNSRPAATHRASAAAGSMTATNSNGNCSREY	10296
Qy	761	GAGCCACACGCGCGCCGTGCATGACGAGACCGGACGCGGAGCGCAGACCCGAAAGCA	820
Db	10297	STRASASRARCGGNSHYSGSRASNGYAGASNGYSASAGNTSASNSPASNAGGHHSTH	10356
Qy	821	GCGGCGCAAGAGGACCGGCGCAAGCCACAGCGGCGCTAGGCGCCTGGG-CCACGCGCGGCA	879
Db	10357	RGAAGTSPAAAHASNTSYAHSHFGGASYSASNSASCTGGGAAYCTGGAGNNGSR	10416
Qy	880	CGGCGGAGAGCGGGCGGAGACCGCAGAGAGCCGACCGGACCGGCGGCGCAAGG	939
Db	10417	MTSRYSGCTSSRYSHASASYSASYSRARAGASVASTRGTHRAAASRGSAAAGYSGNCT	10476
Qy	940	CACGACAGCCACCGCA 954	
Db	10477	HRASNGYSGNTSAA 10491	

## RESULT 8

AF429315 125020 bp DNA linear PRI 18-JAN-2002  
 LOCUS Home sapiens junctophilin 3 (JPH3) gene, partial cds.  
 DEFINITION AF429315  
 ACCESSION AF429315  
 VERSION AF429315.1 GI:17646244  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homini; Homo.  
 REFERENCE 1 (bases 1 to 125020)  
 Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S., Ingersoll-Ashworth, R.G., Fleisher, A., Stevanin, G., Brice, A., Potter, N.T., Ross, C.A. and Margolis, R.L.  
 A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2  
 Nat. Genet. 29 (4), 377-378 (2001)  
 JOURNAL  
 MEDLINE 21563737  
 PUBMED 11694876  
 REFERENCE 2 (bases 1 to 125020)  
 Holmes, S.E., Ingersoll-Ashworth, R.G., Ross, C.A. and Margolis, R.L.  
 Direct Submission  
 Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA  
 JOURNAL  
 TITLE Location/Qualifiers  
 FEATURES  
 source 1..125020  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="16"  
 /map="16q24.3; between D16S520 and WI-12410"  
 /note="Isolated from a patient with Huntington's disease-like 2 (HDL2)"  
 repeat\_region complement(35581..35746)  
 /rpt\_type=tandem  
 /rpt\_unit="ctg"  
 complement(36507..36887)  
 /gene="JPH3"  
 /note="synonym: JF3"  
 complement(36507..36887)  
 /gene="JPH3"  
 /note="component of the junctional complex between plasma membrane and endoplasmic reticulum"  
 /product="junctophilin 3"  
 /protein\_id="AA140941.1"  
 /db\_xref="GI:17646245"  
 /translation="MSSGRFNFDDGSGYCGMEDGKAHGIVCTGPKGGEYTGSMG HGFVYGVYTRMSGNTYQGTWAGKRGHIGLESKKWYVKGEMTHGFGRGVRECAG NGKRYHGTWNSGLQDGYETETSDG"  
 ORIGIN  
 Query Match 5.3%; Score 50.8; DB 9; Length 125020;  
 Best Local Similarity 12.0%; P-Val. No. 0.51;  
 Matches 54; Conservative 200; Mismatches 192; Indels 4; Gaps 1;  
 QY 498 CACACCGGTGTGGCGGCACANTATGATGATGATTCGCCCGGATACAGACCATGGA 557  
 Db 17136 SRSAMCCSYACKMCVSYGMSMKYVYACSRGMSKRCYCMRGSTYSTGCGCCT 17195  
 QY 558 TTTCCCGCGGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617  
 Db 17196 TTTCCCGCGGTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17255  
 QY 618 GACCTAAAGAGTTGACCGCGGACCGGACGAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 677  
 Db 17256 GACCTAAAGAGTTGACCGCGGACCGGACGAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 17315  
 QY 678 CACAGGGCGGACAGCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 737

Db 17316 CMKRCYCMRSMRSMRSMYMYASWKSBSRGYCTRCYCMSSKSCYKSYMMRSMK 17375  
 QY 738 CGAAGAGAGATPACAGAGAGCGGACCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGG 797  
 Db 17376 GAKSRCKMSGMSGASBSRSCYKCYKSYKRCYKSYKRCYKSYKRCYKSYKRCYKSYKRCY 17435  
 QY 798 CGAGGAGCGGACAGCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGG 857  
 Db 17436 RWWGKSAMTRMR 17491  
 QY 858 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 917  
 Db 17492 KASYSYMSYMTCT 17551  
 QY 918 CACGAGCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 947  
 Db 17552 CTSRGAMSCWRCRCYWRGASMSRAGSMR 17581  
 RESULT 9  
 AC142034/c  
 LOCUS  
 DEFINITION Rattus norvegicus clone CH230-253B2, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\* 54 unordered pieces.  
 ACCESSION AC142034  
 VERSION AC142034.1 GI:29135505  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 123570)  
 Wuzny, D., Marie, Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amn, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Bacc, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhammed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Casar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Gre, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Decker, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Davies, K., Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flegel, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gama, R., Garcia, A., Garner, T., Garza, M., Gatregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guetara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyar, C., Kowis, C., Kraft, C.L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Louie, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Mallory, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenan, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundada, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwachukwu, O., Okunolu, G., Olariupunagoon, A., Pal, S., Parke, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindeexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzos, R., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reid, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodery, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savary, G., Scherer, S., Scott, G., Shaleman, S., Shen, H., Shelby, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,  
Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villaseña, D.,  
Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,  
Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R.,  
Wooden, H., Worley, K., Wright, D., Wright, R., Yab, J., Yakub, S.,  
Yeo, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,  
Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,  
Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.  
Direct Submission  
Unpublished  
2 (bases 1 to 123570)  
Morley, K. C.  
Submitted (21-MAR-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, Ore  
Baylor Plaza, Houston, TX 77030, USA  
-----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
-----  
Project Information  
Center project name: GQOT  
Center clone name: CH230-253B2  
-----  
Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 91744 bases at least Q40  
Consensus quality: 100630 bases at least Q30  
Consensus quality: 105715 bases at least Q20  
Estimated insert size: 90869; sum-of-coverage estimation  
Quality coverage: 1x in Q20 bases; sum-of-coverage estimation

-----  
\* NOTE: Bactinated insert size may differ from sequence length  
\* (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html>).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 54 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
1 1419: contig of 1419 bp in length  
\* 1420 1519: gap of unknown length  
\* 1520 2583: contig of 1064 bp in length  
\* 2584 2684: gap of unknown length  
\* 2684 4098: contig of 1415 bp in length  
\* 4099 4198: gap of unknown length  
\* 4199 5410: contig of 1212 bp in length  
\* 5411 5511: gap of unknown length  
\* 5511 6856: contig of 1346 bp in length  
\* 6857 6957: gap of unknown length  
\* 6957 8164: contig of 1208 bp in length  
\* 8164 8265: gap of unknown length  
\* 8265 9302: contig of 1038 bp in length  
\* 9303 9402: gap of unknown length  
\* 9402 10679: contig of 1277 bp in length  
\* 10680 10779: gap of unknown length  
\* 10779 11779: contig of 1000 bp in length  
\* 11780 11879: gap of unknown length  
\* 11880 13188: contig of 1309 bp in length  
\* 13189 13288: gap of unknown length  
\* 13289 14534: contig of 1246 bp in length  
\* 14535 14634: gap of unknown length  
\* 14635 15660: contig of 1026 bp in length  
\* 15661 15760: gap of unknown length  
\* 15761 18008: contig of 2248 bp in length  
\* 18009 18108: gap of unknown length  
\* 18109 19731: contig of 1623 bp in length  
\* 19732 19832: gap of unknown length  
\* 19832 21073: contig of 1242 bp in length

21074 21173: gap of unknown length  
\* 21174 21176 21173: contig of 1002 bp in length  
\* 21176 22275: gap of unknown length  
\* 22276 23779: contig of 1504 bp in length  
\* 23780 23879: gap of unknown length  
\* 23880 25886: contig of 2007 bp in length  
\* 25887 25986: gap of unknown length  
\* 25987 27663: contig of 1777 bp in length  
\* 27664 27863: gap of unknown length  
\* 27864 29028: contig of 1165 bp in length  
\* 29029 29128: gap of unknown length  
\* 29129 30386: contig of 1258 bp in length  
\* 30387 30486: gap of unknown length  
\* 30487 32341: contig of 1555 bp in length  
\* 32342 32441: gap of unknown length  
\* 32442 34065: contig of 1624 bp in length  
\* 34066 34165: gap of unknown length  
\* 34166 35979: contig of 1814 bp in length  
\* 35980 36079: gap of unknown length  
\* 36080 37895: contig of 1816 bp in length  
\* 37896 37995: gap of unknown length  
\* 37996 40144: contig of 2149 bp in length  
\* 40145 40244: gap of unknown length  
\* 40245 42243: contig of 1995 bp in length  
\* 42244 42343: gap of unknown length  
\* 42344 43405: contig of 1062 bp in length  
\* 43406 43505: gap of unknown length  
\* 43506 45570: contig of 2065 bp in length  
\* 45571 45670: gap of unknown length  
\* 45671 47982: contig of 2312 bp in length  
\* 47983 48082: gap of unknown length  
\* 48083 49373: contig of 1231 bp in length  
\* 49374 49473: gap of unknown length  
\* 49474 51351: contig of 1878 bp in length  
\* 51352 51451: gap of unknown length  
\* 51452 53148: contig of 1637 bp in length  
\* 53149 53248: gap of unknown length  
\* 53249 54816: contig of 1568 bp in length  
\* 54817 54916: gap of unknown length  
\* 54917 57844: contig of 2928 bp in length  
\* 57845 57944: gap of unknown length  
\* 57945 60066: contig of 2122 bp in length  
\* 60067 60166: gap of unknown length  
\* 60167 61534: contig of 1368 bp in length  
\* 61535 61634: gap of unknown length  
\* 61635 64403: contig of 2769 bp in length  
\* 64404 64503: gap of unknown length  
\* 64504 67778: contig of 3275 bp in length  
\* 67779 67878: gap of unknown length  
\* 67879 69295: contig of 1417 bp in length  
\* 69296 69395: gap of unknown length  
\* 69396 71592: contig of 2197 bp in length  
\* 71593 71692: gap of unknown length  
\* 71693 74596: contig of 2904 bp in length  
\* 74597 74696: gap of unknown length  
\* 74697 76912: contig of 2216 bp in length  
\* 76913 77012: gap of unknown length  
\* 77013 79045: contig of 2033 bp in length  
\* 79046 79145: gap of unknown length  
\* 79146 82135: contig of 2989 bp in length  
\* 82135 82234: gap of unknown length  
\* 82235 85286: contig of 3052 bp in length  
\* 85287 85386: gap of unknown length  
\* 85387 88347: contig of 2961 bp in length  
\* 88348 88447: gap of unknown length  
\* 88449 92453: contig of 4006 bp in length  
\* 92454 92553: gap of unknown length  
\* 92554 96685: contig of 4132 bp in length  
\* 96686 96785: gap of unknown length  
\* 96786 101131: contig of 4346 bp in length  
\* 101132 101231: gap of unknown length  
\* 101232 106183: contig of 4952 bp in length  
\* 106184 106283: gap of unknown length



```

Db      50718 ABDAGVCRCEMYKACADBCYKRDSCWSAHWYVWYBBHBSKSRBGRHMHACIKSM 50659
Qy      908 GGAGCGGACGACGACCGGACCGGCC 934
Db      50658 CKSVYKVMWVKMKCYWCRCSCWBCS 50632

RESULT 11
AD000671/c
LOCUS   AD000671          46251 bp    DNA        linear    PRI 10-DEC-1997
DEFINITION Homo sapiens DNA from chromosome 19-cosmid f24109 containing HRX2,
            genomic sequence.
ACCESSION AD000671
VERSION   AD000671.1 GI:1905893
KEYWORDS  chromosome 19.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 46251)
            Lamerdin,J.E.
            Direct Submission
            Submitted (27-NOV-1996) J.E. Lamerdin, Human Genome Center,
            Lawrence Livermore National Laboratory, 7000 East Ave, Livermore,
            CA, USA, 94551 jane@cgf.llnl.gov ow@cornak.llnl.gov
2 (bases 1 to 46251)
            Lamerdin,J.E., McCreedy,P.M., Adamson,A.W., Burkhardt-Schultz,K.,
            Garcia,E., Kyle,A., Ramirez,M., Stiiwagen,S., Ganes,J.,
            Dargann,L., Bruce,R., Quan,G., Montgomery,M., Ow,D., Kobayashi,A.,
            Olsen,A.O. and Carrino,A.V.
            Direct Submission
            Submitted (27-NOV-1996) Sequence analysis of a 1 Mb region in human
            19q13.1
            GSD8:S:1010603.
            constructed at LINTL from flow-sorted chromosomes
            from hybrid UVSHL9-5B, which carries chromosome 19 as its only
            human chromosome
            Overlaps cosmid R28052 (by
            9.5 kb) to the left and is
            1.4 kb to the left of cosmid P25965
            map=19q13.1 between D19S208 and CAMPS
            Human Genome Center
            Biology and Biotechnology Research Program
            Lawrence Livermore National Laboratory
            7000 East Avenue
            Livermore, CA 94550 USA.
            Location/Qualifiers
            1..46251
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="19"
            /cell_line="UVSHL9-5B"
            /cell_type="fibroblast"
            /clone_id="LI19NC02 F2 chromosome 19-specific cosmid
            library"
            536..850
            /note="repeat match = HSA104796; putative"
            /rpt_family="Alu"
            1142..1189
            /note="repeat match = HSA113971; putative"
            /rpt_family="Alu"
            1153..1752
            /note="repeat match = HSA113971; putative"
            /rpt_family="Alu"
            /note="repeat match = HSA102794; putative"
            /rpt_family="Alu"
            1484..2226
            /note="repeat match = HSA106431; putative"
            /rpt_family="Alu"
            2108..2257
            /note="repeat match = HSA102931; putative"
            /rpt_family="Alu"
            2561..2609
            /note="repeat match = HSA118610; putative"

FEATURES
source
    1..46251
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="19"
    /cell_line="UVSHL9-5B"
    /cell_type="fibroblast"
    /clone_id="LI19NC02 F2 chromosome 19-specific cosmid
    library"
    536..850
    /note="repeat match = HSA104796; putative"
    /rpt_family="Alu"
    1142..1189
    /note="repeat match = HSA113971; putative"
    /rpt_family="Alu"
    1153..1752
    /note="repeat match = HSA113971; putative"
    /rpt_family="Alu"
    /note="repeat match = HSA102794; putative"
    /rpt_family="Alu"
    1484..2226
    /note="repeat match = HSA106431; putative"
    /rpt_family="Alu"
    2108..2257
    /note="repeat match = HSA102931; putative"
    /rpt_family="Alu"
    2561..2609
    /note="repeat match = HSA118610; putative"

repeat_region
    /rpt_family="Alu"
    2571..2730
    /note="repeat match = HSA106183; putative"
    /rpt_family="Alu"
    complement(2878..3169)
    /note="repeat match = HSA100369; putative"
    /rpt_family="Alu"
    complement(3131..3179)
    /note="repeat match = HSA105507; putative"
    /rpt_family="Alu"
    3221..3500
    /note="repeat match = Alu; putative"
    /rpt_family="Alu"
    complement(4313..4415)
    /note="repeat match = HSA100485; putative"
    /rpt_family="Alu"
    complement(4334..4629)
    /note="repeat match = HSA106577; putative"
    /rpt_family="Alu"
    5110..5729
    /note="repeat match = HSA100284; putative"
    /rpt_family="Alu"
    6002..6035
    /note="repeat match = HSA117280; putative"
    /rpt_family="Alu"
    6011..6133
    /note="repeat match = HSA118162; putative"
    /rpt_family="Alu"
    6038..6271
    /note="repeat match = Alu; putative"
    /rpt_family="Alu"
    6790..7065
    /note="ss region; putative"
    7280..7543
    /note="repeat match = Alu; putative"
    /rpt_family="Alu"
    7475..7563
    /note="repeat match = HSA102791; putative"
    /rpt_family="Alu"
    7695..7936
    /note="repeat match = HSA102508; putative"
    /rpt_family="Alu"
    7847..7960
    /note="repeat match = HSA107958; putative"
    /rpt_family="Alu"
    10138..11022
    /note="predicted exon, gra12exons human_1.3; frame=0,
    forward strand, quality=excellent; putative"
    11736..11852
    /note="similarity: sp|Q05516|PLZF_HUMAN; ZINC FINGER
    PROTEIN PLZF (PROMYELOCYTIC LEUKEMIA ZINC...); similarity:
    gi|913049; mPLZF(B)=promyelocytic leukemia zinc finger
    protein; putative"
    11988..12077
    /note="similarity: gi|545177; PLZF=ZINC FINGER
    protein(retinoic acid receptor alpha, RAR alpha...);
    similarity: sp|Q05516|PLZF_HUMAN; ZINC FINGER PROTEIN PLZF
    (PROMYELOCYTIC LEUKEMIA ZINC...); similarity: gi|913049;
    mPLZF(B)=promyelocytic leukemia zinc finger protein;
    putative"
    12608..12748
    /note="predicted exon, gra12exons human_1.3; frame=1,
    forward strand, quality=good; putative"
    13530..13892
    /note="predicted exon, gra12exons human_1.3; frame=2,
    forward strand, quality=good; putative"
    15698..15771
    /note="predicted exon, gra12exons human_1.3; frame=0,
    forward strand, quality=good; putative"
    16358..16532
    /note="predicted exon, gra12exons human_1.3; frame=2,
    forward strand, quality=good; putative"
    16422..16502
    /note="predicted exon, gra12exons human_1.3; frame=2,
    forward strand, quality=good; putative"

```

```

/misc_feature /note="similarity: pit|B48205; All-1 protein -GTE form -
mouse (fragment); similarity: |D14540; (D14540) MLL [Homo
sapiens]; similarity: |L22179; (L22179) MLL-AF4 der(11)
fusion protein [Homo sapiens]; putative"
16656.16703
/misc_feature /note="similarity: |D14540; (D14540) MLL [Homo sapiens];
similarity: |L22179; (L22179) MLL-AF4 der(11) fusion
protein [Homo sapiens]; similarity: pit|B48205; All-1
protein -GTE form - mouse (fragment); putative"
16806.16865
/misc_feature /note="similarity: pit|B48205; All-1 protein -GTE form -
mouse (fragment); similarity: |L22179; (L22179) MLL-AF4
der(11) fusion protein [Homo sapiens]; similarity:
|D14540; (D14540) MLL [Homo sapiens]; putative"
17224.17315
/misc_feature /note="predicted exon, grai12exons_human_1.3; frame=2,
forward strand, quality=excellent; putative"
17305.17454
/misc_feature /note="ss region; putative"
17305.17454
repeat_region complement(17521..17812)
/note="repeat match = HSA106523; putative"
/misc_feature /rpt_family="Alu"
17843.17983
/misc_feature /note="predicted exon, grai12exons_human_1.3; frame=1,
forward strand, quality=excellent; putative"
17998.18065
/misc_feature /note="ss region; putative"
18079.18229
/misc_feature /note="predicted exon, grai12exons_human_1.3; frame=0,
forward strand, quality=excellent; putative"
18088.18254
/misc_feature /note="similarity: |D14540; (D14540) MLL [Homo sapiens];
similarity: pit|A44265; trithorax-homologous protein
(translocation) HRX - human; similarity: |L22179;
(L22179) MLL-AF4 der(11) fusion protein [Homo sapiens];
similarity: |L04284; (L04284) HRX [Homo sapiens];
similarity: |D14540; (D14540) MLL [Homo sapiens];
putative"
18088.18222
/misc_feature /note="similarity: |L22179; (L22179) MLL-AF4 der(11)
fusion protein [Homo sapiens]; putative"
18661.18786
/misc_feature /note="similarity: pit|A44265; trithorax-homologous
protein (translocation) HRX - human; similarity: |L22179;
(L22179) MLL-AF4 der(11) fusion protein [Homo sapiens];
similarity: |L04284; (L04284) HRX [Homo sapiens];
similarity: |D14540; (D14540) MLL [Homo sapiens];
putative"
18664.18785
/misc_feature /note="predicted exon, grai12exons_human_1.3; frame=0,
forward strand, quality=good; putative"
18958.19027
/misc_feature /note="predicted exon, grai12exons_human_1.3; frame=1,
forward strand, quality=excellent; putative"
19371.19517
/misc_feature /note="predicted exon, grai12exons_human_1.3; frame=1,
forward strand, quality=excellent; putative"
19579.19626
/misc_feature /note="predicted exon, grai12exons_human_1.3; frame=0,
forward strand, quality=good; putative"
19727.19775
repeat_region /note="repeat match = HSA103439; putative"
19739.20004
repeat_region /rpt_family="Alu"
/note="repeat match = HSA101336; putative"
Query Match 5.2%; Score 50.2; DB 9; Length 46251;
Best Local Similarity 47.6%; Pred. No. 0.74; Indels 0; Gaps 0;
Matches 146; Conservative 0; Mismatches 163;

```

```

Db 13559 GCACATCCGCCGCCGCCGCCGCCGCCCATCTTGACACCGGTAGAGGGGCCGGAGGCG 13500
Qy 752 CGAGGACGCGAGCCACACGCGGCCCGCGTGATGAGACGAGACGCGCGAGGACGCCACA 811
Db 13499 GGGGGAGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13440
Qy 812 CCGGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 871
Db 13439 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 13380
Qy 872 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 931
Db 13379 ACCGCGAGGCGCACACGCGGAGAGGAGATCGCGCGCGCGCGCGCGCGCGCGCGCG 13320
Qy 932 GCCACGGCGCGAG 942
Db 13319 GGGGAGCGACAG 13309

RESULT 12
AC022315 194835 bp DNA linear HTG 25-APR-2002
LOCUS Homo sapiens chromosome 19 clone RP11-38C1, WORKING DRAFT SEQUENCE,
DEFINITION 18 unordered pieces.
AC022315
AC022315 GI:17975245
VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULFILL.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 194835)
Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,
FedeSPIEL,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,
Mao,J., Komp,C., Kotler,S., Lam,B., Marathe,R., Miranda,M.,
Morehouse,A.J., Nguyen,M., Oefner,P., Palm,C.J., Ramirez,D.,
Southwick,A.M., Webb,C., Wilhelmy,J., Yu,S. and Davis,R.W.
Unpublished
2 (bases 1 to 194835)
Bruno,D., Conn,L., Dela Rosa,M., FedeSPIEL,N., Foreman,P.,
Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,
Morehouse,A.J., Oefner,P., Palm,C.J., Ramirez,D.,
Wilhelmy,J., Yu,S. and Davis,R.W.
Direct Submission
Submitted (30-JAN-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
3 (bases 1 to 194835)
Bruno,D., Conn,L., Dela Rosa,M., FedeSPIEL,N., Foreman,P.,
Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,
Morehouse,A.J., Oefner,P., Palm,C.J., Ramirez,D.,
Wilhelmy,J., Yu,S. and Davis,R.W.
Direct Submission
Submitted (25-APR-2002) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
On Dec 21, 2001 this sequence version replaced gi:17063116.
----- Genome Center
Center: Stanford DNA Sequencing and Technology Development
Center
Center code: SDSFDC
Web site: http://sequence-www.stanford.edu/group/human/
Contact: hum-info@sequence.stanford.edu
----- Project Information
Center project name: RP11-38C1
Center clone name: RP11-38C1
----- Summary Statistics
Sequencing Vector: M13mp18; X02513; 96% of reads
Sequencing Vector: plasmid; plasmid accession; 4% of reads
Chemistry: Dye-primer; 12% of reads
Chemistry: Dye-terminator Big Dye; 87% of reads
Assembly program: Phrap; version 0.990319

```



Consensus quality: 184441 bases at least Q40  
 Consensus quality: 186605 bases at least Q30  
 Consensus quality: 187666 bases at least Q20  
 Insert size: 168558; agarose-fp  
 Insert size: 193135; sum-of-coverage  
 Quality coverage: 14.4x in Q20 bases; agarose-fp  
 Quality coverage: 12.6x in Q20 bases; sum-of-coverage  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 18 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 1604: contig of 1604 bp in length  
 \* 1605 1700: gap of unknown length  
 \* 1705 3770: contig of 2066 bp in length  
 \* 3771 3870: gap of unknown length  
 \* 3871 5909: contig of 2039 bp in length  
 \* 5910 6009: gap of unknown length  
 \* 6010 8095: contig of 2086 bp in length  
 \* 8096 8195: gap of unknown length  
 \* 8196 10705: contig of 2510 bp in length  
 \* 10706 10805: gap of unknown length  
 \* 10806 12952: contig of 2147 bp in length  
 \* 12953 13052: gap of unknown length  
 \* 13053 16988: contig of 3936 bp in length  
 \* 16989 17088: gap of unknown length  
 \* 17089 20048: contig of 2860 bp in length  
 \* 20049 20149: gap of unknown length  
 \* 20149 23005: contig of 2857 bp in length  
 \* 23006 23105: gap of unknown length  
 \* 23106 27314: contig of 4209 bp in length  
 \* 27315 27415: gap of unknown length  
 \* 27415 31443: contig of 4029 bp in length  
 \* 31443 31543: gap of unknown length  
 \* 31544 37789: contig of 6246 bp in length  
 \* 37790 37889: gap of unknown length  
 \* 37890 44274: contig of 6385 bp in length  
 \* 44275 44374: gap of unknown length  
 \* 44375 53077: contig of 8703 bp in length  
 \* 53078 53177: gap of unknown length  
 \* 53178 60886: contig of 7709 bp in length  
 \* 60887 60986: gap of unknown length  
 \* 60987 77446: contig of 16460 bp in length  
 \* 77447 77546: gap of unknown length  
 \* 77547 107542: contig of 29996 bp in length  
 \* 107543 107643: gap of unknown length  
 \* 107643 194835: contig of 87193 bp in length.

## FEATURES

## source

1. 194835  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="RP11-38C1"  
 /clone\_lib="RP11 human BAC library 11"  
 1. 1604  
 /note="assembly\_name:Contig53"  
 misc\_feature  
 1705. 3770  
 /note="assembly\_name:Contig54"  
 misc\_feature  
 3871. 5909  
 /note="assembly\_name:Contig58"  
 misc\_feature  
 6010. 8095  
 /note="assembly\_name:Contig50  
 clone\_end:77"  
 misc\_feature  
 8196. 10705  
 /note="assembly\_name:Contig51"  
 misc\_feature  
 10806. 12952  
 /note="assembly\_name:Contig52"  
 misc\_feature  
 13053. 16988  
 /note="assembly\_name:Contig53"

misc\_feature 17089. 20048  
 /note="assembly\_name:Contig54"  
 misc\_feature 20149. 23005  
 /note="assembly\_name:Contig55"  
 misc\_feature 23106. 27314  
 /note="assembly\_name:Contig56"  
 misc\_feature 27415. 31443  
 /note="assembly\_name:Contig57"  
 misc\_feature 31544. 37789  
 /note="assembly\_name:Contig58"  
 misc\_feature 37890. 44274  
 /note="assembly\_name:Contig59"  
 misc\_feature 44375. 53077  
 /note="assembly\_name:Contig60"  
 misc\_feature 53178. 60886  
 /note="assembly\_name:Contig61"  
 misc\_feature 60987. 77446  
 /note="assembly\_name:Contig62"  
 misc\_feature 77547. 107542  
 /note="assembly\_name:Contig63  
 clone\_end:SP6"  
 misc\_feature 107643. 194835  
 /note="assembly\_name:Contig64"

## ORIGIN

Query Match 5.2%; Score 50.2; DB 2; Length 194835;  
 Best Local Similarity 47.6%; Pred. No. 0.67;  
 Matches 148; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 632 GACCGCGCAGCGCGCCGCAAGAACGAGCTTTCAGACGAAATCCAAACAGAGGCGCCGACA 691  
 DB 184206 GCGCCCG 184265  
 QY 692 CCGCGACGCGACCG 751  
 DB 184266 GCAACTGCG 184325  
 QY 752 CGAGACGCGAGCGCACCG 811  
 DB 184326 GGGGGAGGGGGCG 184385  
 QY 812 CCGGAGCGACG 871  
 DB 184386 GCG 184445  
 QY 872 GCG 931  
 DB 184446 AGCGCGAGGCGCACCGGAGAGGAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 184505  
 QY 932 GCCCGAGCGAG 942  
 DB 184506 GGGGAGGACGAG 184516

## RESULT 13

AC143344/C 152686 bp DNA linear HTG 10-APR-2003  
 LOCUS AC143344  
 DEFINITION Macaca mulatta clone CH250-268P23, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
 AC143344  
 AC143344.1 GI:29611668  
 VERSION  
 KEYWORDS HTG; HTGS PHASE2; HTGS\_PG1.  
 SOURCE Macaca mulatta (rhesus monkey)  
 ORGANISM Macaca mulatta  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.  
 1 (bases 1 to 152686)  
 Ceausu, M. and Milosavljevic, A.  
 Pooled genomic indexing (Pgi): mathematical analysis and experiment  
 design  
 (in: Guigo R. and Gusfield D. (Eds.):  
 ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI  
 2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-26;

## REFERENCE

AUTHORS  
 TITLE  
 JOURNAL

## REFERENCE AUTHORS

Springer (2002)

2 (bases 1 to 152688)  
Milosavljevic A., Sodre

Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Direct Submission  
Unpublished

3 (bases 1 to 152686,  
Worley & C

Direct Subm

Submitted (08-APR)

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 152686)

Worley, K.C.

Direct Submitted (10-APR)

of Molecular and Human Genetics, Baylor College of Medicine, One

COMMENT

Center: Baylor College of Medicine  
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

----- Project Information -----  
 contact: ngbcs-help@bcm.tmc.edu

Center project name: LDMM

Center clone name: CH250-268P23

-----  
Summary statistics  
Chemistry: Dye-primer Bodipy: T

Chemistry: Dye-terminator Big Dye

Consensus quality: 3754 bases at least Q40  
Consensus quality: 4014 bases at least Q30

consensus quality: 4914 bases at least Q20  
consensus quality: 5932 bases at least Q20

[illegible]

## FEATURES

```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)
* NOTE: The contigs are based on the application
* of the Pei method using the Human genome (NCBI build 31)
* as the comparative genome.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 152686: contig of 152666 bp in length.
* Location/Qualifiers
* 1..152686

```

ORIGIN

Query Match	5.2%;	Score 50;	DB 2;	Length 152686;
Best Local Similarity	48.9%;	Pred. No. 0.76;		
Matches 131;	Conservative	0;	Mismatches 137;	Indels 0;
			Gaps	0;

[illegible]

RESULT 14  
ACCEPTED

LOCUS	AK655393	2000 bp	DNA	linear	PAT 22-MAR-2003
DEFINITION	Sequence 5263 from Patent WO03000089.				
ACCESSION	AK655393				
VERSION	AK655393.1	GI:29158207			
KEYWORDS					

**SOURCE**  
**ORGANISM**

REFERENCE  
1  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

## NOTES

FEATURES	location/Qualifiers
JOURNAL	Syngenta Participations AG (CH)
TITLE	Plant genes involved in defense against pathogens
Patent:	WO 03000898-A 5263 03-JAN-2003;

```

/db_xref="taxon:4530"
ORIGIN

```

## ORIGIN



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 14:04:56 ; Search time 364.27 Seconds  
(Without alignments)  
1160.755 Million cell updates/sec

Title: US-10-076-747-7

Perfect score: 957

Sequence: 1 ttggcgttactaatgcatg.....ggcagcagagccacgcagct 957

Scoring table: IDENTITY NUC

Searched: 3373663 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : N\_Geneseq\_2972n04:\*

1: geneseq1360s:\*\n2: geneseq1390s:\*\n3: geneseq2000s:\*\n4: geneseq2001as:\*\n5: geneseq2002s:\*\n6: geneseq2003as:\*\n7: geneseq2003bs:\*\n8: geneseq2003cs:\*\n9: geneseq2004s:\*\n10: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	956	99.9	957	ABX92193	ABX92193 Human ova
2	809.6	84.6	1460	ABX92194	ABX92194 Human ova
3	68.2	7.1	2067	ABA09659	ABA09659 Human bon
4	55.2	5.8	349	ABT07668	ABT07668 Human bre
5	55	5.7	1190	AA883947	AA883947 DNA encod
6	51	5.3	908	AA887524	AA887524 DNA encod
7	50.8	5.2	10732	AA810594	AA810594 Gene enco
8	49.6	5.2	2000	ADA71938	ADA71938 Rice gene
9	48.4	5.1	114955	AA53491	AA53491 Human ade
10	44.4	4.6	1451	AA502527	AA502527 Human col
11	44.2	4.6	87980	AA53223	AA53223 (5 of
12	43	4.5	366	ABX43944	ABX43944 Bovine BS
13	43	4.5	764	ADA10928	ADA10928 Human CDN
14	43	4.5	6225	AA55273	AA55273 Human enz
15	43	4.5	6225	AA54721	AA54721 Human enz
16	43	4.5	6225	AA54721	AA54721 Human enz
17	43	4.5	6225	AA54721	AA54721 Human enz
18	43	4.5	6225	AA54721	AA54721 Human enz
19	42.8	4.5	35459	AB296537	AB296537 Human nuc
20	42.4	4.4	5760	AB297130	AB297130 Human nuc
21	42.4	4.4	5059	AA584332	AA584332 Stealyot
22	42.4	4.4	5617	AA57163	AA57163 Partial h
23	42.4	4.4	8091	AA57001	AA57001 Human Not

C 24	42.4	4.4	8091	6	AB234974	AB234974 Human gen
C 25	42.4	4.4	8257	6	AB570408	AB570408 Human bon
C 26	42.4	4.4	168174	6	ABT11173	ABT11173 Human 5-1
C 27	42.4	4.4	168173	6	ABT11114	ABT11114 Human 5-1
C 28	42.2	4.4	895	6	ABV83635	ABV83635 Human bre
C 29	42.2	4.4	4705	7	ACA61566	ACA61566 Streptomy
C 30	42.2	4.4	35384	3	AAF21436	AAF21436 Human enz
C 31	42.2	4.4	109519	5	AA506693	AA506693 Micromono
C 32	42	4.4	691	6	ABQ40803	ABQ40803 Oligonuc
C 33	42	4.4	756	6	ABQ40802	ABQ40802 Oligonuc
C 34	42	4.4	38064	6	ABA99456	ABA99456 Actinopla
C 35	42	4.4	712	6	ABQ40859	ABQ40859 Actinopla
C 36	41.8	4.4	712	6	ABQ40858	ABQ40858 Oligonuc
C 37	41.8	4.4	4488	2	AA555520	AA555520 Human thr
C 38	41.6	4.3	4570	9	AA557372	AA557372 Human thr
C 39	41.4	4.3	37314	4	AAK71358	AAK71358 Human imm
C 40	41.4	4.3	172570	6	ABQ88207	ABQ88207 Human ost
C 41	41.4	4.3	920	4	AAH21491	AAH21491 Barley al
C 42	41.2	4.3	1164	4	AA551496	AA551496 Pseudomon
C 43	41.2	4.3	1164	7	ACA19500	ACA19500 Prokaryot
C 44	41.2	4.3	154746	6	AA255519	AA255519 Human her

## ALIGNMENTS

RESULT 1	ABX92193	ABX92193 standard; CDNA, 957 BP.
ID	ABX92193	ABX92193
AC	ABX92193	ABX92193
XX	ABX92193	ABX92193
DT	08-MAY-2003	(first entry)
DE	Human ovarian specific nucleic acid DEX0310_7.	
XX	Human, gene; ss; ovarian specific nucleic acid; ONSA; ovarian cancer;	
KW	non-cancerous ovarian disease; gene therapy; vaccine; cyostatic;	
KX	gynaecological.	
XX		
OS	Homo sapiens.	
XX		
PN	WO20022785-A2.	
PD	21-NOV-2002.	
XX		
PF	13-FEB-2002; 2002WO-US022271.	
XX		
PR	13-FEB-2001; 2001US-0268290P.	
XX		
PR	15-FEB-2001; 2001US-026834P.	
XX		
XX	(DIAD-) DIADEXUS INC.	
XX	Saleeda S, Macina RA, Hu P, Recipon H, Karra K, Caffery R;	
XX	Sun Y, Liu C;	
XX	WPI; 2003-120677/11.	
XX	P-PSDB; AB061022.	
PT	New isolated ONSA nucleic acid and encoded polypeptide, useful for	
PT	identifying, diagnosing, monitoring, staging, imaging and treating	
PT	ovarian cancer and non-cancerous diseases in ovarian tissues.	
PS	Claim 1; Page 145; 224pp; English.	
XX		
XX	The invention relates to a new isolated nucleic acid termed ovarian	
XX	specific nucleic acid (ONSA) comprising: (a) a nucleic acid molecule that	
XX	encodes any of 53 fully defined protein sequences appearing as AB061018-	
XX	AB061070 (termed ovarian specific proteins, OSP); (b) any of 76 fully	
XX	defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a	
XX	sequence having at least 60% sequence identity to the nucleic acid	
XX	molecule of (a) or (b). Also included are a method for determining the	



```

DB 634 AAGCTACAGCTGCCAGGTGAC-----GCATGAAGGAG-CACCGTGATGAAGACGT 685
QY 183 GGTCCCCCTACAGAAATGTTCTATAGTTCTCTAACACGCTGAGCCCCCTACCAATGCG 242
DB 686 GGTCCCCCTACAGAAATGTTCTATAGTTCTCTAACACGCTGAGCCCCCTACCAATGCG 745
QY 243 GAGACTAGCAGCTGCGAGGATGCCAAGGGGAGAGGGGCTCTCCATCCACACGACCAAG 302
DB 746 GAGACTAGCAGCTGCGAGGATGCCAAGGGGAGAGGGGCTCTCCATCCACACGACCAAG 805
QY 303 GGGGAGTCAAAAGCCCTTATCATCGCGATGTCGACGTATGTAAAGGCGCTCAAAATTA 362
DB 806 GGGGAGTCAAAAGCCCTTATCATCGCGATGTCGACGTATGTAAAGGCGCTCAAAATTA 865
QY 363 GATATTCCTGCACTTGTTGAAATGCTCTACATATATAAAGACATATCTCACTACA 422
DB 866 GATATTCCTGCACTTGTTGAAATGCTCTACATATATAAAGACATATCTCACTACA 925
QY 423 CGGAGTCGATGATCACCGGTCCTGCGGGGGAATGACACTTCGCTCTGCGGTCCCAAT 482
DB 926 CGGAGTCGATGATCACCGGTCCTGCGGGGGAATGACACTTCGCTCTGCGGTCCCAAT 985
QY 483 GATGACTCATATGTTACACACCGGTGTGGGGCGACANCTATGATGATATTTGCCCGGG 542
DB 986 GATGACTCATATGTTACACACCGGTGTGGGGCGACANCTATGATGATATTTGCCCGGG 1045
QY 543 ATCAACAGCATGATGATTTCCCGGTGGCTGATGAAATATGCGGTACCGGGCATCAAAAT 602
DB 1046 ATCAACAGCATGATGATTTCCCGGTGGCTGATGAAATATGCGGTACCGGGCATCAAAAT 1105
QY 603 CGCCCGGAGGTACAGACTTAAAAAGTTGAACCGCGGACGCGCGGAAGAACAGGCTTTG 662
DB 1106 CGCCCGGAGGTACAGACTTAAAAAGTTGAACCGCGGACGCGCGGAAGAACAGGCTTTG 1165
QY 663 ACGGAATGCCAACAACAGAGGGCGCGACCGGAGAGCGACCCCGGGCGAGAGAGCCCA 722
DB 1166 ACGGAATGCCAACAACAGAGGGCGCGACCGGAGAGCGACCCCGGGCGAGAGAGCCCA 1225
QY 723 CTGGCGGCGAGGCGAGGCGAAGACAGATACGAGGACCGGACACCGCGCCCGCTGCA 782
DB 1226 CTGGCGGCGAGGCGAGGCGAAGACAGATACGAGGACCGGACACCGCGCCCGCTGCA 1285
QY 783 TGAAGCGAGACAGCGGCCGAGGAGCGCGACGCCGAGACCGCGCTCAAGACCGCGCA 842
DB 1286 TGAAGCGAGACAGCGGCCGAGGAGCGCGACGCCGAGACCGCGCTCAAGACCGCGCA 1345
QY 843 AGCCACACGCGCTTAGGCGCTGCGCCACGCGCGGCGACGCGCGAGAGGCGGCGAGCA 902
DB 1346 AGCCACACGCGCTTAGGCGCTGCGCCACGCGCGGCGACGCGCGAGAGGCGGCGAGCA 1405
QY 903 CCGCAGAGAGACCGACACGAGACCGGACCGGCGCAAGGCGAGAGGACCGGAGCT 957
DB 1406 CCGCAGAGAGACCGACACGAGACCGGACCGGCGCAAGGCGAGAGGACCGGAGCT 1460

```

RESULT 3  
ABA09659/c  
ID ABA09659 standard; DNA; 2067 BP.

ABA09659;  
15-JAN-2002 (first entry)

Human bone marrow expressed oligonucleotide SEQ ID NO: 168.

Human; bone marrow; cytostatic; antirheumatic; antiarthritic; vulnery;  
antiflammatory; antibacterial; immunosuppressive; vasotropic; cancer;  
antiparkinsonian; neuroprotective; nootropic; haemostatic; osteopathic;  
antituber; fungicide; antidiabetic; antiaesthetic; antiallergic;  
immunostimulant; analgesic; cerebroprotective; antianaemic; infection;  
nervous system disorder; autoimmune disorder; inflammation; allergy; ds.

```

OS Homo sapiens.
XX WO200174836-A1.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US010472.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX 30-NOV-2000; 2000US-02505839.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Dmanac RT, Ford JE, Boyle BJ;
XX WPI, 2001-626375/72.
XX New bone marrow-expressed nucleic acids and polypeptides, useful for
XX diagnosis, treatment of inflammatory, autoimmune, neurological, cancer
XX remodeling.
XX Claim 1, Page 248-249; 380pp; English.
XX The present invention relates to bone marrow expressed polynucleotides
XX and proteins. These sequences can be used in the treatment of
XX inflammatory conditions (eg arthritis, Crohn's disease), cancer, central
XX and peripheral nervous system diseases and neuropathies, such as
XX Alzheimer's, Parkinson's and Huntington's diseases, spinal cord
XX disorders, head trauma, cerebrovascular diseases, myeloid and lymphoid
XX cell disorders, platelet disorders, stem cell disorders, bone
XX degenerative disorders, autoimmune disorders, for example multiple
XX sclerosis, diabetes and arthritis, viral and bacterial infections,
XX allergies and blood coagulation disorders. The present sequence is a DNA
XX of the invention
XX Sequence 2067 BP; 418 A; 579 C; 620 G; 450 T; 0 U; 0 Other;
XX
XX Query Match 7.1%; Score 68.2; DB 5; Length 2067;
XX Best Local Similarity 63.6%; Pred. No. 6.6e-08;
XX Matches 138; Conservative 0; Mismatches 73; Indels 6; Gaps 2;
QY 79 GCTATCTAGAGCTGACGCGCTGATGAGTGAAGTCCACAGAGCTACAGCTGAC 138
DB 987 GCTATCTAGAGCTGACGCGCTGATGAGTGAAGTCCACAGAGCTACAGCTGAC 930
QY 139 AGAGTCTTAGCTGATGAAGGAGACACCGTGAATGAAGAGAGTGTCCCTACAGAAAT 198
DB 929 -CAGGTCACGCGCATTTGAAGGGAGACACCGTGGAGAGAGACAGGGCCCCCTACAGCAAT 871
QY 199 GTTCTATAGTTCCTTAAACAGCTCAGCCCCCACTACCAATGGCGAGACTAGACGCTGC 258
DB 870 GTTCTATAGTTCCTTAAACAGCTCAGCCCCCACTACCAATGGCGAGACTAGACGCTGC 814
QY 259 AGGATCCCAAGGAGAGGGGCTCTCCATCCACACC 295
DB 813 GGAATCCCAAGGAGAGGGGCTCTCTCTCCACCC 777

```

RESULT 4  
ABT07668  
ID ABT07668 standard; cDNA; 349 BP.

ABT07668;  
14-NOV-2002 (first entry)

Human breast cancer associated coding sequence SEQ ID NO: 147.

Human; breast specific gene; breast specific protein; breast cancer;  
gene therapy; cytostatic; gene; ss.

```

OS Homo sapiens.
XX
XX PN MO200264611-A1.
XX
XX PD 22-AUG-2002.
XX
XX PF 12-FEB-2002; 2002WO-US004197.
XX
XX PR 13-FEB-2001; 2001US-0268292P.
XX
XX PA (DIAD-) DIADEXUS INC.
XX
XX PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R,
XX PI Sun Y, Liu C;
XX DR WPI; 2002-657582/70.
XX
XX PT New breast specific nucleic acids and proteins, useful for identifying,
XX PT diagnosing, monitoring, staging, imaging, and treating breast cancer and
XX PT non-cancerous disease states in breast tissue, and in gene therapy.
XX
XX PS Claim 1; Page 266; 367pp; English.
XX
XX CC The present invention provides human breast specific coding sequences and
XX CC proteins. These can be used in the diagnosis and treatment of breast
XX CC cancer. non-cancerous diseases of the breast. The present sequence is
XX CC a coding sequence of the invention
XX
SQ Sequence 349 BP; 109 A; 64 C; 81 G; 95 T; 0 U; 0 Other;

Query Match 5.8%; Score 55.2; DB 6; Length 349;
Best Local Similarity 82.9%; Pred. No. 0.00011;
Matches 63; Conservative 0; Mismatches 13; Indels 0; Gaps 0.

QY 4 GGCTTACTATCATCATGCTCGACGCGCCGCAAGTGTGATGATGCTGGTCGCGGCCGA 63
Db 26 GGTTCATCTTAATGACATGCTCGACGCGCGCCGCAAGTGTGATGATGCTGGTCGCGGCCA 85
QY 64 GGTAACGGGCTCAGCAG 79
Db 86 GGTCACAGTTTAGCTG 101

RESULT 5
AAS83947
ID AAS83947 standard; cDNA; 1190 BP.
XX
XX AC AAS83947;
XX
XX DT 13-FEB-2002 (first entry)
XX
XX DE DNA encoding novel human diagnostic protein #19751.
XX
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US008631.
XX
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00645167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YF,
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDS; ABG19760.

```

XX	New isolated polynucleotide and encoded polypeptides; useful in
PT	diagnostics; forensics; gene mapping; identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
PX	Claim 1; SEQ ID NO 19751; 103bp; English.
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (II) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64157-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: the sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
SQ	Sequence 1190 BP; 277 A; 348 C; 288 G; 263 T; 0 U; 14 Other;
	Query Match 5.7%; Score 55; DB 5; Length 1190;
	Best Local Similarity 60.4%; Pred. No. 0.00018;
	Matches 131; Conservative 0; Mismatches 75; Indels 11; Gaps 2;
QY	79 GCTATCTGAGCGGTACAGCGCCTTGATGCAGTGGAGAGTCCCAAGAAGTACAGCCTGCAC 138
Db	677 GCTATCTGAGCGGTACAGCGCCTTGAGCAGTGGAGATTCCACAGAAAGCTACAGCTGCCAG 736
Db	139 AGAAGTCTTAAGTCGATGAAGGAGACAACCGTGATGAAGACATGTGTCCTTACAGAAAT 198
Db	737 GG-----TCAAGCATTTAAGAGGGAGCACCGGTGGAGAAAGACAGTGGCCCTTACAGATG 791
QY	139 GTTCTATAAGTTCCTCTAACACAGCTCACGCCCCCACTACCAAGGGAGACTGACAGCTGCC 258
Db	792 TTCATATAGTTTCTCAACCCCTTCAACCCCCCACCACAGGAGATGACTTAGA-----GCTGC 845
QY	259 AGGATCCCAAGAGAGAGGGTCTCTTCATCCACACC 295
Db	846 AGATTTCCACAGGAGAGGGGTCTTCTCTCCACACC 882
RESULT 6	
ID	AAS87524
XX	AAS87524 standard; cDNA; 308 BP.
XX	
AC	AAS87524;
XX	
DT	13-FEB-2002 (first entry)
DE	DNA encoding novel human diagnostic protein #23328.
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
PD	11-OCT-2001.
PP	30-MAR-2001; 2001WO-US008631.

XX 21-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
XX MPI: 2001-639362/73.  
DR P-PSDB; ABG23337.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnosis, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 1; SEQ ID NO 23328; 103bp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 908 BP; 213 A; 290 C; 228 G; 177 T; 0 U; 0 Other;  
Query Match 5.3%; Score 51; DB 5; Length 908;  
Best Local Similarity 68.5%; Pred. No. 0.0019;  
Matches 146; Conservative 0; Mismatches 55; Indels 12; Gaps 5;  
QY 74 CAGCAGCTCTGAGCCTGACGCGCTGATGAGTGAAGATCCACAGAGTACAGCC 133  
DB 623 CAGCAGCTCTGAGCCTGAGCCTGA-GCAGTGGGA-AGTCCACAGAGCTACAGCT 680  
QY 134 TGCACAGACTCTTACCTGATGAGAGGAGACACCGTGTATGAAGACAGTGTCCCTTACA 193  
DB 681 TGC---CAGGGTCACGGCAGCATGAGAGGAGCACCGTGGAGAGACAGTGGCCCTTACA 737  
QY 194 GAATGTTCTATAGGTCTCTCAACGCTCAAGCCCCCACTACCAATGCGAGACTAGAC 253  
DB 738 G-AAAGTTCATAGGTTTCTCAACCCCTCACTCCCCCCCA-----CGGAAACTTAG 790  
QY 254 GCTGACGAGATCCCAAGGAGAGGGGCTCTCC 286  
DB 791 ACTTCAGAAATCCACGGGAGGGGTTTCTCTCC 823  
RESULT 7  
AAA10594  
ID AAA10594 standard; DNA; 10732 BP.  
XX  
XX AAA10594;  
AC  
XX  
XX 29-JUN-2000 (first entry)  
DT  
XX  
DE Gene encoding a subunit of cellulose synthase.

KW Cellulose synthase; cellulose production; increase yield; ds.  
XX  
XX Vigna angularis.  
OS  
XX JP2000060568-A.  
PN  
XX  
XX 29-FEB-2000.  
PD  
XX  
XX 26-AUG-1998; 98UP-00239998.  
PF  
XX  
XX 26-AUG-1998; 98UP-00239998.  
PR  
XX  
XX 26-AUG-1998; 98UP-00239998.  
PA (MIZU/) MIZUNO K.  
PA (OJIP) OJI PAPER CO.  
XX  
XX MPI: 2000-342371/30.  
DR P-PSDB; AAY85179.  
XX  
XX A gene encoding a cellulose synthetic equipment - for the improvement in  
PT the amount of cellulose synthesized in a plant body.  
PT  
XX  
XX Claim 2; Page 14-21; 32pp; Japanese.  
PS  
XX  
XX This sequence represents a gene encoding a subunit of the cellulose  
CC synthase complex of Vigna angularis. The invention relates to subunits of  
CC cellulose synthetic equipment, that can be used to increase the amount of  
CC cellulose synthesised by a plant. The proteins and genes encoding them  
CC can also be used to improve the properties of the cellulose being  
CC produced by a plant  
XX  
SQ Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 0 U; 2251 Other;  
Query Match 5.3%; Score 50.8; DB 3; Length 10732;  
Best Local Similarity 14.8%; Pred. No. 0.0047;  
Matches 135; Conservative 333; Mismatches 445; Indels 2; Gaps 2;  
QY 41 GATGATGCTGCTGCGCGCGGAGTACCGGCTGACAGCTATCTGAGCCTGACGCTG 100  
DB 9578 GASHSHARGSGRGYSYSGRSRYSRVAGYAAASRAASAMNTASASVAGMTGSAAV 9637  
QY 101 ATGCAGTGTAGAGTCCACAGAGCTACAGAGCTGACAGAGTCTAGCTGATGAAGG 160  
DB 9638 AVASRVAGYSAATHRGNSRSRAASRYSAAVATHRGASRSRGTHRGYSGSRTTR 9697  
QY 161 AGACACCGTGTGAAAGACAGTGTCCCTACAGAAATGTTCTAATGTTCTTAACAG 220  
DB 9698 RGGGYSGRYAASHSRSGTHRTGTRHRTGVAVAASMTGRTSRASSRSRY 9757  
QY 221 CTGACGCCCCACATCAAGATGCGAGACTGACACGCTGACGAGATCCCAAGGAGAGG 280  
DB 9758 SSRYASNTHTCTSSRAHSYSGTTHRVANGYTRGNAVAVASTHGNASHSVATHEH 9817  
QY 281 CTCTCATCCACACACACAGAGGCGAGTCAAAAGCCCTTATCATCGGATGTCAGCTC 340  
DB 9818 SYSGYASASHASYSVAAAARHSHVAAASASHASASRSRARGYSRYVAHBSATH 9877  
QY 341 ATGTAAAGGCGCTCAAAATTAATATTCTGCACTGTGTAATCTCTCATATATA 400  
DB 9878 RGNASNRNGCYSHSYSSRYSSRSASRYVAASNSRRAAGVAGASVAAASRSYASG 9937  
QY 401 CAAGAACAATCACTACACAGAGATGATGATCACCAGCTGCGGCGAGATGCC 460  
DB 9938 GRSRNGAS-SRYSRGYSGVAGAGRAAVAASTYRASRSRAGNSGAGYARAGAH 9996  
QY 461 ACTTCGCTCTGCTGCTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 520  
DB 9997 SRTYRSASNSGYWTRTHRYTMTVAAGVAGHNTRYSTHRSRGSRYVAGHSHVAC 10056  
QY 521 TATGAGTGAATATGCGCGGAGTCAAGACATGATGATGATGATGATGATGATGATGAT 580  
DB 10057 STYRASRGYSTRASRYTRGTAASNAAGYTRTHRAAASMTMTASNHSCYSHSRNABV 10116  
QY 581 ATGGGATGACGCGGATCAAAATTCGCGGAGCTACAGACTAAAAAAGTGAACGCGCA 640



```

Db      10117 AARSMNTARGAATAYSNGNHASTPAGSTHRARGSRASGYGASNGTTHRHVAIGSGARG 10176
QY      641 GCGGCCGAAGAAACAGGCTTTCAGACGAATGCGAAACACAGAGGCGCCGACGCCGACAGC 700
Db      10177 HSRARGASNSRSRGNVATHTHRHSRCYSGHSAAGNASNGYATRYTASNGYSGHTH 10236
QY      701 GACCCCGGCGGAGAGAGCCCACTGCGGACGAGCGAGGAGGAGACAGATACGAGAGCC 760
Db      10237 RASYSASYSRAGSAGSYGAGARGISGNSRAATHTASARGGSRMTASNGNGSBRGY 10296
QY      761 GAGCCACACGCGCGCCCTGCATGACACGAGACGCGCGAGCGCGACGCCGAAACA 820
Db      10297 STHRASRARGGNSRHYSRPSASNGYNGASNGYSASNGYSASNSRASNAGSNGHTH 10356
QY      821 GCGCGCGAAGAGCGACCGCCGCAACGACGCGCCTAGCCCTGGG-CCACGCGCGGCGCA 879
Db      10357 RGAAGTSAALAAHSASNSYVAHSRGGASYSASNSRASNCTYSGGAAATCYSGGNGNGSR 10416
QY      880 CCGCGAGAGCGCGGCGGAGACCGACGAGAGACCGACCGACCGCGCGCGCGCAAGG 939
Db      10417 MHSRYGCGYSRSHASASYSASYSRARGASNASTRGTHRAAASRGYSAAAGCYSGNGT 10476
QY      940 CAGCAGAGCCACCGCA 954
Db      10477 HRASNGYSGNYSAA 10491

RESULT 8
ADA71938
ID      ADA71938 standard; DNA; 2000 BP.
XX
XX      ADA71938;
XX
DT      20-NOV-2003 (first entry)
XX
DE      Rice gene, SEQ ID 5263.
XX
KW      Plant; bacterial infection; fungal infection; viral infection; rice;
KW      gene; de.
XX
XX      Oryza sativa.
XX
OS      WO2003000898-A1.
XX
PN      03-JAN-2003.
XX
PD      22-JUN-2001; 2001WO-1B001105.
XX
PF      22-JUN-2001; 2001WO-1B001105.
XX
PR      22-JUN-2001; 2001WO-1B001105.
XX
PA      (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI      Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI      Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR      MPI; 2003-175290/17.
XX
PT      Identifying at least one gene involved in plant resistance or response to
PT      pathogenic infection for conferring resistance or tolerance to a plant to
PT      bacterial, fungal or viral infection by determining or detecting plant
PT      gene expression.
XX
PS      Claim 27; SEQ ID NO 5263; 899bp; English.
XX
XX      The present invention relates to a method (M1) for identifying genes
XX      involved in plant resistance or response to pathogenic infection. M1
XX      comprises identifying a gene whose expression is significantly altered in
XX      the incompatible interaction of plant gene expression relative to
XX      expression of the gene in an uninfected plant; in a mutant plant that
XX      does not express a gene associated with response to pathogenic infection,
XX      or in a corresponding incompatible or compatible interaction. (M1) is
XX      useful for conferring resistance to resistance or tolerance to a plant to

```

```

CC      bacterial, fungal or viral infection. The present sequence was used to
CC      illustrate the invention.
XX
SQ      Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
XX
Query Match      5.2%; Score 49.6; DB 7; Length 2000;
Best Local Similarity 8.7%; Pred. No. 0.0058;
Matches 40; Conservative 217; Mismatches 202; Indels 0; Gaps 0;

QY      491 ATAGTTACACACCGGTGTGTCGCGGCACANCTATGATGATATTGCCCCGGGATACACA 550
Db      75 RTGSKSSGSSYSQKMKKKYKSKRWGRGRGRMRSPRRWGRYRCARSGMAGSGSR 134
QY      551 CCATGATTTCCCGCGTGCATGATGATGCGGTAACGGCATCAATTCGCGCGG 610
Db      135 MMGKSRMYWYMCYARGGSCCKRKSCKGSGMKTCTRRARSGSGMSAKTKSGMSGR 194
QY      611 AGCTACAGACTTAAATAATTGACCGCGAGCGCGCGGAGAAACAGGCTTTCAGACGATG 670
Db      195 MMSSCGRSGGRSAYSRVGTSTRKYGTYKQWYTSABRCMRAYWTTYSWACSSYTW 254
QY      671 CCAAACAGAGGCGCGGACCGGACCGGACCGGACCGGCGGAGAGAGCCCACTGCGGCA 730
Db      255 RSKRSMWMMKMKWMSRYSYMSYSYMMCTAKYSYRWCYTYGGGWRGATRW 314
QY      731 GCGCGAGCGGGAAGACAGATACGAGACCGGACGACCGCGCGCGCGCTGCATGAGACG 790
Db      315 GREGYMRMAMTYKYWYGYGKMGKMGWAGMMWRMSRCWKGACYWRMWRMTRRR 374
QY      791 AGACGCGCGAGAGGAGCGGACCGGACGAGACGCGCGCGGAGAGACCGCGGACGACG 850
Db      375 WAKKSRTRRKRKMKRKYKMRGYSRMRSCRRAMRMRCSRGRMKXGCRGCMTC 434
QY      851 GCGCCTAGGCGCTGCGCCACGCGCGCGGCGGACGCGGAGAGCGGCGGAGACCGGAGA 910
Db      435 RMKSYGMRMKSKRMAKSKYKMSRMYRKKKCSRTTMGTGRGMMGTMRCHYKRS 494
QY      911 GACGACGACGAGCGGACCGGCGCGGCGGAGAGAGCC 949
Db      495 GMRKCRRRMRGMYRMRKRYMSARYMTARCARKYS 533

RESULT 9
AAK53491/C
ID      AAK53491 standard; DNA; 114955 BP.
XX
XX      AAK53491;
XX
DT      05-JUL-1999 (first entry)
XX
DE      Human adenosine A1 receptor antisense oligonucleotide fragment.
XX
XX      Antisense oligonucleotide; multiple target; antisense treatment;
XX      impaired respiration; inflammation; lung disease;
XX      pulmonary vasoconstriction; inflammation; allergic rhinitis;
XX      acute asthma; allergy; asthma; impeded respiration;
XX      respiratory distress syndrome; pain; cystic fibrosis;
XX      pulmonary hypertension; pulmonary vasoconstriction; emphysema;
XX      chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;
XX      colon cancer; breast cancer; lung cancer; pancreatic cancer;
XX      hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;
XX      prostate cancer; ss.
XX
OS      Synthetic.
XX
XX      WO9913886-A1.
XX
PN      25-MAR-1999.
XX
PD      17-SEP-1998; 98WO-US019419.
XX
PF      17-SEP-1997; 97US-0059160P.
XX
PR      09-JUN-1998; 98US-00093972.

```

XX (UVEC-) UNIV EAST CAROLINA.  
 PA Nyce JW;  
 XX WPI; 1999-229400/19.  
 XX  
 PT New antisense oligonucleotides used in treatment of, e.g. pulmonary  
 PT vasoconstriction.  
 XX  
 PS Disclosure; Page 37; 120pp; English.  
 XX  
 CC The specification describes antisense oligonucleotides (AA52869-X55271)  
 CC directed against at least 2 mRNAs selected from target genes, coding and  
 CC non-coding regions of RNAs corresponding to target genes, gene initiation  
 CC codons, genomic flanking regions, intron-exon borders, the 5' end, the 3'  
 CC end and the junction between coding and non-coding regions and all  
 CC segments of RNAs encoding proteins associated with one or more diseases,  
 CC conditions or mixtures. The antisense oligonucleotides may be derived  
 CC from sequences AA55272-74. These multiple target oligonucleotides  
 CC (specifically AA55180-271) can be used for the antisense treatment of  
 CC diseases and conditions. Typical diseases and conditions are those  
 CC associated with impaired respiration and inflammation, including lung  
 CC diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,  
 CC acute asthma, allergies, asthma, impeded respiration, respiratory  
 CC pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary  
 CC disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.  
 CC colon cancer, breast cancer, lung cancer, pancreatic cancer,  
 CC hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as  
 CC well as all types of cancers which may metastasize or have metastasized  
 CC to the lungs, including breast and prostate cancer  
 XX  
 SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 0 U; 21427 Other;  
 Query Match 5.1%; Score 48.4; DB 2; Length 114955;  
 Best Local Similarity 42.6%; Pred. No. 0.042;  
 Matches 109; Conservative 9; Mismatches 138; Indels 0; Gaps 0;  
 QY 678 CAGAGGCGCCGACAGCCGCGAGCGACCCCGCGAGAGCCCACTGCGCGCAAGCGAG 737  
 Db 107080 CANNHNNSCGGGCGGGCGGGGGAAGCCGGCGCGGGCCGAGCCAGNNHNNNSCCGG 107021  
 QY 738 GCGAAGACACATATCGAGACCGGAGCCACACCGCGCCCGCTCATGAGACGAGCGCG 797  
 Db 107020 GCGCGGGGCGAGCA 106961  
 QY 798 GGAAGGAGCGGAGACCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTA 857  
 Db 106960 AGCGGGGCGCGGGCGCGAGCCAGGNNHNNNSCCGGCGCGCGCGCGCGCGCGCG 106901  
 QY 858 GGCCTTGCAGCGCGCGCGCGCGCGAGCGCGCGCGCGAGCCGACGAGACCGAGC 917  
 Db 106900 GGGGCGGAGCGAGGNNHNNNSCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 106841  
 QY 918 CAGCGACCGGAGCGCGC 933  
 Db 106840 CAGGCGCANNHNNNSCC 106825  
 RESULT 10  
 ID AA02527 standard; cDNA; 1451 BP.  
 XX AA02527;  
 AC  
 DT 19-MAY-2000 (first entry)  
 XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:2518.  
 XX Human colon cancer; tumour; diagnosis; gene expression product; probe;  
 KM detection; cancerous state; metastasis; identification; breast cancer;  
 KM oestrogen receptor-positive breast cancer; therapy;

KM oestrogen receptor-negative breast cancer; lung cancer; ss.  
 XX Homo sapiens.  
 OS  
 XX WO958675-A2.  
 PN  
 XX 18-NOV-1999.  
 XX  
 PD 13-MAY-1999; 99MO-US010602.  
 XX  
 PR 14-MAY-1999; 98US-0085426P.  
 PR 15-MAY-1998; 98US-008537P.  
 PR 15-MAY-1998; 98US-0085696P.  
 PR 21-OCT-1998; 98US-0105234P.  
 PR 27-OCT-1998; 98US-0105877P.  
 XX  
 PA (CHR) CHIRON CORP.  
 PA (HSE-) HYSEQ INC.  
 XX  
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
 PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;  
 PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;  
 PI Leskowitz D, Kita D, Garcia V, Jones LM, Staehle-Crain B;  
 XX WPI; 2000-126369/11.  
 DR  
 XX Polynucleotide library used to determine cancerous states of mammalian  
 XX cells.  
 PS Claim 1; Page 1015-1016; 1097pp; English.  
 XX  
 CC AAA00010 to AAA02716 represent polynucleotides isolated from cDNA  
 CC libraries constructed from human colon cancer cell lines. The present  
 CC invention also describes a method of detecting differentially expressed  
 CC genes correlated with a cancerous state of a mammalian cell, comprising  
 CC detecting at least one differentially expressed gene product in a test  
 CC sample derived from a cell suspected of being cancerous, where detection  
 CC of the differentially expressed gene product is correlated with a  
 CC cancerous state of the cell from which the test sample was derived. The  
 CC polynucleotide sequences can be used in a method for detecting  
 CC differentially expressed genes correlated with a cancerous state of a  
 CC mammalian cell. The polynucleotides can also be used as probes for  
 CC detecting and mapping related genes. They can be used in diagnosis and  
 CC prognosis of diseases and disorders (e.g. identification of pre-  
 CC metastatic or metastatic cancerous states, stages of cancer, or  
 CC responsiveness of cancer to therapy). This is particularly for breast  
 CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-  
 CC negative breast cancer, lung cancer, and colon cancer  
 XX  
 SQ Sequence 1451 BP; 167 A; 363 C; 551 G; 23 T; 0 U; 347 Other;  
 Query Match 4.6%; Score 44.4; DB 3; Length 1451;  
 Best Local Similarity 37.8%; Pred. No. 0.13;  
 Matches 108; Conservative 0; Mismatches 178; Indels 0; Gaps 0;  
 QY 650 GAACAGGCTTTCAGACGAAATCCAAACACAGAGGCGCGGACCGGCGGACCCCGG 709  
 Db 316 GACGNGGCGNCCGANNCGANNCGANNCGANNCGANNCGANNCGAACAACNCCCGGNG 375  
 QY 710 CGAGAGACCCCACTGCGCGAGGCGGAGCGGAGGACGATATCGAGAGCGGAGCCAC 769  
 Db 376 GGGNNANCNCCGGGGCGCGGNNANCCAGNANNCGNAGGAGGCGCCCGCGGCGNNNGN 435  
 QY 770 GCGCGCCCTGTCATGAGACGAGACGGCGGAGGAGCGGAGACCGCGAGCGCGCGCA 829  
 Db 436 GCGNNGGANNNGGAGCGCGGNNNGAGAGCGGCGGAGGCGGAGCGCGCGCGCGCG 495  
 QY 830 GAGCGACCGCGCAAGCCACACGCGCTTAGCGCTTGCAGCGCGCGCGCGCGCGAG 889  
 Db 496 GCGAGNNCCNCCGCGGNNCCCGCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 555  
 QY 890 GCGGCGGAGACACCGGAGGAGACCGGAGACCGGAGACCGGAGCGCGCGCGCGCGCG 935

XX  
PR 11-OCT-2000: 2000US-0239841P

XX (LASEX) LASEX A W.  
 PA (JONE) JONES D A.  
 XX Lasek AW, Jones DA;  
 XX WPI; 2003-265756/26.  
 DR New combination comprising cDNAs that are differentially expressed in  
 PT colon disorder, useful for diagnosing, treating, staging or monitoring  
 PT treatment for colon cancers.  
 PS Claim 1; SEQ ID NO 46; 231bp; English.  
 XX The invention relates to a combination comprising cDNAs that are  
 CC differentially expressed in colon disorder. The methods and compositions  
 CC of the present invention are useful for diagnosing, treating, staging or  
 CC monitoring treatment for colon cancer. They are also useful in high  
 CC throughput methods for using cDNAs to detect differential expression of  
 CC nucleic acids in a sample, screening molecules or compounds to identify a  
 CC ligand which specifically binds a cDNA and using a protein to screen  
 CC molecules or compounds to identify at least one ligand which specifically  
 CC binds the protein. The present sequence represents a human cDNA  
 CC differentially expressed in colon cancer.  
 CC  
 SQ Sequence 764 BP; 168 A; 229 C; 214 G; 152 T; 0 U; 1 Other;  
 Query Match 4.5%; Score 43; DB 8; Length 764;  
 Best Local Similarity 75.0%; Pred. No. 0.24; Mismatches 10; Gaps 4;  
 Matches 108; Conservative 0; Mismatches 26; Indels 10; Gaps 4;  
 QY 69 GCGGTGACGAGCTATCTGAGCGCTGATGAGTGTGAGAGTCCACAGAACTA 128  
 Db 620 GGGGCGACGAGCTATCTGAGCGCTGAC-GCCTGAGCCAGTGGGA-AGTCCACAGAACTA 677  
 QY 129 CAGGCTGACGAGCTATCTGAGCTGATGAGGAGACCCCTGATG--AAGAAGTGTGTC 186  
 Db 678 CAGGTGCCA-----GGTCANGGATGAGAGGGAGCACCTGTGAGAGAGACAGTGGCC 731  
 QY 187 CCCTACAGAAATGTTCTATAGTT 210  
 Db 732 CCTTACAAGATGTTTCTATAGTT 755  
 RESULT 14  
 AAX55273/c  
 ID AAX55273 standard; DNA; 6225 BP.  
 AC AAX55273;  
 XX  
 DT 05-JUL-1999 (first entry)  
 DE  
 XX Human enzyme-related antisense polynucleotide.  
 XX Antisense oligonucleotide; multiple target; antisense treatment;  
 KM impaired respiration; inflammation; lung disease;  
 KM pulmonary vasoconstriction; inflammation; allergic rhinitis;  
 KM acute asthma; allergy; aschma; impeded respiration;  
 KM respiratory distress syndrome; pain; cystic fibrosis;  
 KM pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
 KM chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;  
 KM colon cancer; breast cancer; lung cancer; pancreatic cancer;  
 KM hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
 KM prostate cancer; ss.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX MO9913886-A1.  
 XX 25-MAR-1999.  
 PD 17-SEP-1998; 98WC-US019419.  
 XX

XX 17-SEP-1997; 97US-0058160P.  
 PR 09-JUN-1998; 98US-00093972.  
 XX (UYEC-) UNIV EAST CAROLINA.  
 PA  
 XX Myce JW;  
 XX WPI; 1999-229400/19.  
 DR  
 XX New antisense oligonucleotides used in treatment of, e.g. pulmonary  
 PT vasoconstriction.  
 PT  
 PS Disclosure; Page 24-25; 120bp; English.  
 XX The specification describes antisense oligonucleotides (AAX52869-X55271)  
 CC directed against at least 2 mRNAs selected from target genes, coding and  
 CC non-coding regions of RNAs corresponding to target genes, gene initiation  
 CC codons, genomic flanking regions, intron-exon borders, the 5'-end, the 3'-  
 CC end and the juxta-section between coding and non-coding regions and all  
 CC segments of RNAs encoding proteins associated with one or more diseases,  
 CC conditions or mixtures. The antisense oligonucleotides may be derived  
 CC from sequences AAX55272-74. These multiple target oligonucleotides  
 CC (specifically AAX55180-271) can be used for the antisense treatment of  
 CC diseases and conditions. Typical diseases and conditions are those  
 CC associated with impaired respiration and inflammation, including lung  
 CC diseases, pulmonary vasoconstriction, inflammation, allergic rhinitis,  
 CC acute asthma, allergies, aschma, impeded respiration, respiratory  
 CC distress syndrome, pain, cystic fibrosis, pulmonary hypertension,  
 CC pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary  
 CC disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g.  
 CC colon cancer, breast cancer, lung cancer, pancreatic cancer,  
 CC hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as  
 CC well as all types of cancers which may metastasize or have metastasized  
 CC to the lungs, including breast and prostate cancer  
 CC  
 SQ Sequence 6225 BP; 13 A; 2111 C; 2185 G; 1580 T; 0 U; 336 Other;  
 Query Match 4.5%; Score 43; DB 2; Length 6225;  
 Best Local Similarity 42.9%; Pred. No. 0.47;  
 Matches 135; Conservative 20; Mismatches 160; Indels 0; Gaps 0;  
 QY 639 CAGCGCGCGAAGACAGGCTTCCAGCGAATGCCAATACAGAGGCGCCGACCGGACG 698  
 Db 2652 CAGCAGVCVAVAGCGCGCAGVCACGCGGAGGAGACAGCGGAGAAACCCACVC 2593  
 QY 699 GCGACCCCGGCGGAGAGAGCCCACTGCGGAGGCGGAGGAGAAAGACAGATACAGAGAC 758  
 Db 2592 CCAGVCCACCCVGGGCGGCGGCGGAGVCAVCAVGAAGACAGAGAGACCAAGGCC 2533  
 QY 759 GCGAGCCACACGCGCGCGCTGTCATGAGACGAGACGCGCCGAGGAGCGCAGACCCGAG 818  
 Db 2532 CCGGCGCCGACGGAAGCGCGGAAAGAAAGACCGGACCGCAGCGCGCGCCACCCAC 2473  
 QY 819 CAGCGCGCGAAGAGAGACCGGCGCAAGCAGCGGCTTATGCGCGCGCGCGCGCGCC 878  
 Db 2472 CGGCCACCCGAGAGCGACACACGACGCTTCACTGAAGAGCGAGCGCGCGAGAGACCA 2413  
 QY 879 ACGCCGAGAGAGCGGCGGAGACCGCGAGAGACCGACCGAGACCCGCGCGCGAG 938  
 Db 2412 GGAGAGCGGCGCGCGAGCGGCGGAGCGGCGGAGCGGCGGCGGAGCGGAGCGGAG 2353  
 QY 939 GCGAGCGAGCGGACCG 953  
 Db 2352 AAGACGAGGAGVACV 2338  
 RESULT 15  
 AAA34721/c  
 ID AAA34721 standard; DNA; 6225 BP.  
 AC AAA34721;  
 XX

DT 28-JUL-2000 (first entry)  
XX Human adenosine receptor related polynucleotide SEQ ID NO:2410.  
DE  
XX Human, adenosine receptor; low adenosine antisense oligonucleotide;  
XX phosphodiesterase; impaired respiration; inflammation; allergy;  
XX allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
XX antiallergic; antispasmodic; cytoskeletal; analgesic; impaired airway;  
XX lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
XX respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
XX pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
XX cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
XX Homo sapiens.  
XX MO200009525-A2.  
XX 24-FEB-2000.  
XX 03-AUG-1999; 99WO-US017712.  
XX 03-AUG-1998; 98US-0095212P.  
XX (UYEC-) UNIV EAST CAROLINA.  
XX Myce JW;  
XX WPI; 2000-205971/18.  
XX  
XX New antisense oligonucleotides useful for treating e.g. pulmonary  
PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
PT cancers.  
XX  
XX Disclosure; Page 568-569; 1343pp; English.  
XX  
XX The present invention describes a new composition comprising an antisense  
CC oligonucleotide (ON) with low adenosine (up to 15%), which targets  
CC nucleic acids involved in bronchoconstriction, allergies, and/or  
CC inflammation. The ON can have antiinflammatory, antiallergic,  
CC antispasmodic, cytoskeletal and analgesic activities. The compositions are  
CC useful for the treatment of diseases associated with inflammation,  
CC impaired airways, including lung disease and diseases whose secondary  
CC effects afflict the lungs of a subject. They can be used for treating  
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,  
CC impaired respiration, respiratory distress syndrome, pain, cystic  
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,  
CC carcinomas, and cancers which may metastasise to the lungs, including  
CC breast and prostate cancer. The reduction of the adenosine content of the  
CC ONs reduces side effects. The A-containing ONs break down with the  
CC release of deoxyadenosine which activates adenosine receptors causing the  
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the  
CC nucleotide sequences given in the sequence listing from the present  
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185  
CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ  
CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to  
CC AAA33932) are specifically claimed ONs from the present invention. N.B.  
CC Sequences given in the disclosure of the present invention do not match  
CC up with their corresponding SEQ ID NO: sequences given in the sequence  
XX listing  
XX  
XX Sequence 6225 BP; 13 A; 2111 C; 2186 G; 1580 T; 0 U; 335 Other;

Query Match 4.5%; Score 43; DB 3; Length 6225;  
Best Local Similarity 42.9%; Fred. No. 0.47;  
Matches 135; Conservative 20; Mismatches 160; Indels 0; Gaps 0;

DB 2592 CCAGVCCACCCVGGCGGCTCCCGCGVACVAGAAAGACGAGGACCAAGGCC 2533  
QY 759 GCGAGCCACACCGCGCGCCCTGCTATGACGGAAGAGCCCGGAGACCGACAG 818  
DB 2532 CCGGCGCGACCGAGCGCGGAAAGAGCGGAGCCACCGCGCGCCACAC 2473  
QY 819 CAGCGCGCAAGAGCGAGCGCGCAAGCCACACGCGCTAGCCCTGCGCCACGCGCGGCC 878  
DB 2472 CGGCCACCGCGAGCCACACACCGAGGTACGAGAGAGAGCGGACAGAGAGCA 2413  
QY 879 ACAGCGAGAGAGCGCGCGGAGACCGCGAGAGACCGACCGAGCCGCGCCAG 938  
DB 2412 GCAGACGGGAGCCAGCGVCGCGVCAACACAGCCCGGCGCACACCGAGVACAVG 2353  
QY 939 GCAGCAGAGCCACCG 953  
DB 2352 AAGGACGAGGAGVAGC 2338

Search completed: April 26, 2004, 16:12:02  
Job time: 369.27 secs

QY 639 CAGCGCGCGAAGACGCTTTGACGAGATGCCAAACACAGAGCGCGCAGACCGGAG 698  
DB 2652 CAGCAGVCAVAGCGCGVAGCGVAGCGCGGAGAGAGCGCGGAGAGAAACCACTC 2593  
QY 699 GCGACCCCGGCGGAGAGCCCACTGCGGCGAGGCGGAGGCGAAGACAGATACAGAGAC 758



Best Local Similarity 1.9%; Pred. No. 2.6e-05;  
Matches 7; Conservative 219; Mismatches 141; Indels 0; Gaps 0;

```
QY 581 ATCGGTACCGCGCATCAATTGCCCCGGAGCTACAGACTAAAAATTGACCGCGCA 640
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1443 ATTTGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1384
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 641 GCGGCGAAGAACAGGCTTCAGCGAATGCGCAACACAGAGGGCGGACCGCGAGGC 700
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1383 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1324
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 701 GACCCCGGCGGAGAGACCCCACTGCGGAGGCGGAGGCGAAGACAGATACGAGACGC 760
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1323 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1264
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 761 GAGCCACAGCGCGCGCGCGCATGAGACGAGAACCGCGGAGGCGGACCGCAAGCA 820
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1263 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1204
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 821 GCGCGCAAGACGACCGCGGACGACAGCGGCTAGCCCTGCGCCACGCGCGCAC 880
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1203 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1144
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 881 GCGCGAAGGCGCGCGGAGACCGCCAGAGACCAACAGACCGGCGCGCGCGC 940
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1143 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1084
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 941 AGCAGAG 947
    : : : : :
Db 1083 RRRRRRR 1077
```

RESULT 2  
US-09-252-991A-1329  
; Sequence 1329, Application US/09252991A  
; Patent No. 6551735

```
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1329  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-1329
```

Query Match 4.5%; Score 42.8; DB 4; Length 906;  
Best Local Similarity 47.7%; Pred. No. 0.015;

Matches 125; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```
QY 687 GCAGACCGGCGAGCGACCCCGGCGGAGAGACCCCTGCGGCGAGGCGGAGCGAAGAC 746
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 384 GCGGCGCAACAGGAGGTCAAGCGCGCGGACATCGGCTCGATGCGGAGAAACACCCC 443
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 747 AGATCGAGGAGCGGAGCGACACAGCGGCGCGGTGCTATGAGCGGAGACCGGCGAGGAGC 806
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 444 GCCCGAGTACAGGCGCATGAGCCAGAGAGCGCGGCGGACCGCGATGCCCGGCGAGCCGTGC 503
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 807 GCAGACCGGCGAGCGGCGCGCAAGAGCGACCGCGCAAGCGCACAGCGCTTAGGCCCTGCG 866
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 504 CGAGCGCGCGGCGGCGGCGGCGGCTGAGAGAGCGGCGGCGGCGGCGGAGGCGGAG 563
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 867 CCACGCGCGCGCGGCGGAGAGCGGCGGCGGAGCGACCGGAGAGACCGGCGGAGCGC 926
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 564 GAGCAGGCGGATGCGCGAAGAGGTGAAGGCGCTCGGCGTACGCGGCGGCGGAACTTGCGGCC 623
```

```
QY 927 GACCGCGCGGCGGCGGAGCGACAGC 948
    |||||
Db 624 GAGGAATCCAGTCTCCAGAGC 645
```

RESULT 3  
US-09-252-991A-1283  
; Sequence 1283, Application US/09252991A  
; Patent No. 6551735

```
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1283  
; LENGTH: 1023  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-1283
```

Query Match 4.5%; Score 42.8; DB 4; Length 1023;  
Best Local Similarity 47.7%; Pred. No. 0.015;  
Matches 125; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

```
QY 687 GCAGACCGGCGAGCGACCCCGGCGGAGAGACCCCACTGCGGCGAGGCGGAGCGAAGAC 746
    ||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 622 GCGGCGCAACAGGAGGTCAAGCGCGCGGCGGACATGCGTGTGATGGGAGGAAACACCCC 681
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 747 AGATCGAGGAGCGGAGCGACACAGCGCGCGCTGCTATGAGCGGAGAGCGCGGAGGAGC 806
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 682 GCGCGAGTACAGGCGCATGAGCCAGAGCGCGCGGCGGACCGGCGGACATGCCCGGCGAGCGG 741
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 807 GCAGACCGGCGAGCGGCGCGCAAGAGCGGCGCGGAGCGCACAGCGGCTAGGCGCTCGG 866
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 742 CGAGCGCGCGGCGGCGGCGGCGGCTGAGCGAGGCGGAGGCGGCGGAGGCGGCGGAG 801
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 867 CCACGCGCGCGCGGCGGCGGCGGAGGCGGCGGCGGAGCGGCGGAGCGGAGCGGAGCGC 926
    : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 802 GAGCAGGCGGAGTGGCGAGGAGGTGAAGGCGCTCGGCGTACGCGGCGGCGGAACTTGCGGCC 861
    : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 927 GACCGCGCGGCGGCGGAGCGC 948
    |||||
Db 862 GAGGAATCCAGTCTCCAGAGC 883
```

RESULT 4  
US-09-252-991A-1187/c

; Sequence 1187, Application US/09252991A  
; Patent No. 6551735

```
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1187  
; LENGTH: 1080  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa
```

US-09-252-991A-1187

Query Match 4.5%; Score 42.8; DB 4; Length 1080;  
Best Local Similarity 47.7%; Pred. No. 0.016;  
Matches 125; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 687 GGNACCGGCGAGCCACCCCGGCGAGAGCCCTGCGCGAGGCGGAGAGAC 746  
DB 513 GCGGACCAACAGAGGAGTCAAGCCCGCGACATGCTGCTGATGCGAGAAACACCC 454  
QY 747 AGATACGAGGAGCGCGACACCGCGCGCTGTGATAGACGAGAGCGCGAGAGAC 806  
DB 453 GCGGAGTACAGGCGCATGACACGAGCGCGGCGACCGCGACATGCGCGAGCCGTGC 394  
QY 807 GCGAGCCCGGAGCGCGCGCGCAAGACGACCGCGCGACCGCGCTGAGCGCG 866  
DB 393 CGAGCCCGGCGCGAGCGCGCGCTGAGCAGGCGCGCGCGCGAGCGCGCGAG 334  
QY 867 CCGAGCGCGCGCGCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 926  
DB 333 GAGCGAGCGGATGCGCGAGAGGTGAAGGCGCTGCGCGTACGCGAGCGCGAATTGCGCG 274  
QY 927 GACCGCGCGAGCGCGAGAGC 948  
DB 273 GAGGATCCAGTCTCCAGAGC 252

## RESULT 5

US-09-230-652-1/c  
Sequence 1, Application US/09230652A  
Patent No. 6537775

GENERAL INFORMATION:  
APPLICANT: Tournier-Lasserre, Elisabeth  
APPLICANT: Joutel, Anne  
APPLICANT: Bousseer, Marie-Germaine  
APPLICANT: Bach, Jean-Francois  
TITLE OF INVENTION: GENE INVOLVED IN CADASIL, METHOD OF DIAGNOSIS AND  
FILE REFERENCE: 03715.0048-00000  
CURRENT APPLICATION NUMBER: US/09/230.652A  
CURRENT FILING DATE: 1999-05-17  
EARLIER APPLICATION NUMBER: FR 96 09733  
EARLIER FILING DATE: 1996-08-01  
EARLIER APPLICATION NUMBER: FR 97 04680  
EARLIER FILING DATE: 1997-04-16  
EARLIER APPLICATION NUMBER: PCT/FR97/01433  
EARLIER FILING DATE: 1997-07-31  
NUMBER OF SEQ ID NOS: 163  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 8091  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (79)..(7041)  
OTHER INFORMATION: human ADNC No. 6537775ch 3  
US-09-230-652-1

Query Match 4.4%; Score 42.4; DB 4; Length 8091;  
Best Local Similarity 47.7%; Pred. No. 0.045;  
Matches 124; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 675 ACACAGAGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 734  
DB 4112 AGACAGGCG 4053  
QY 735 GAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 794  
DB 4052 CGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3993  
QY 795 GCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 854

DB 3992 TGGCATGAGGAGCGCGCACCGCGGACCTGCACTCCCGCAGAGAGCGCGCACCGCTCGCA 3933  
QY 855 CTAGGCGCTGCG 914  
DB 3932 CGCGGAGCG 3873  
QY 915 GACCGAGGAGCG 934  
DB 3872 GGGCTAGGAGCG 3853

## RESULT 6

US-09-484-970B-65/c  
Sequence 65, Application US/09484970B  
Patent No. 6426186

GENERAL INFORMATION:  
APPLICANT: Jones, Karen A.  
APPLICANT: Volkmuth, Wayne  
APPLICANT: Walker, Michael G.  
TITLE OF INVENTION: BONE REMODELING GENES  
FILE REFERENCE: PB-0014 US  
CURRENT APPLICATION NUMBER: US/09/484.970B  
CURRENT FILING DATE: 2000-01-18  
NUMBER OF SEQ ID NOS: 172  
SOFTWARE: PERL Program  
SEQ ID NO 65  
LENGTH: 8257  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. 6426186 285840.2  
NAME/KEY: unsure  
LOCATION: 7391, 8247  
OTHER INFORMATION: a, t, c, g, or other  
US-09-484-970B-65

Query Match 4.4%; Score 42.4; DB 4; Length 8257;  
Best Local Similarity 47.7%; Pred. No. 0.045;  
Matches 124; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

QY 675 ACACAGAGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 734  
DB 4112 AGACAGGCG 4053  
QY 735 GAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 794  
DB 4052 CGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3993  
QY 795 GCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 854  
DB 3992 TGGCATGAGGAGCGCGCACCGCGGACCTGCACTCCCGCAGAGAGCGCGCACCGCTCGCA 914  
QY 855 CTAGGCGCTGCG 934  
DB 3932 CGCGGAGCG 3873  
QY 915 GACCGAGGAGCG 934  
DB 3872 GGGCTAGGAGCG 3853

## RESULT 7

US-09-962-665-5  
Sequence 5, Application US/09962665  
Patent No. 6537759

GENERAL INFORMATION:  
APPLICANT: Stanton, Jr., Vincent P.  
TITLE OF INVENTION: POLY(POLYGLUTAMATE SYNTHETASE GENE SEQUENCE  
TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE  
FILE REFERENCE: 11926-015004  
CURRENT APPLICATION NUMBER: US/09/962.665



```

1 CURRENT FILING DATE: 2001-09-24
2 PRIOR APPLICATION NUMBER: 09/658,659
3 PRIOR FILING DATE: 2000-09-08
4 PRIOR APPLICATION NUMBER: 09/596,033
5 PRIOR FILING DATE: 2000-06-15
6 PRIOR APPLICATION NUMBER: 09/357,743
7 PRIOR FILING DATE: 1999-07-20
8 PRIOR APPLICATION NUMBER: 09/357,024
9 PRIOR FILING DATE: 1999-07-19
10 PRIOR APPLICATION NUMBER: 60/093,484
11 PRIOR FILING DATE: 1998-07-20
12 NUMBER OF SEQ ID NOS: 16
13 SOFTWARE: FASTSQ For Windows Version 4.0
14 SEQ ID NO 5
15 LENGTH: 3772
16 TYPE: DNA
17 ORGANISM: Homo sapiens
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: 431..441
21 OTHER INFORMATION: n = a or g
22 NAME/KEY: misc_feature
23 LOCATION: 498
24 OTHER INFORMATION: n = c or t
25 NAME/KEY: misc_feature
26 LOCATION: 579..599
27 OTHER INFORMATION: n = g or c
28 US-09-962-665-5

```

Query Match	4.4%	Score 41.8;	DB 4;	Length 3772;
Best Local Similarity	50.8%;	Pred. No. 0.049;		
Matches 153; Conservative	0;	Mismatches 142;	Indels 6;	Gaps 2;

QY	666	TCACCGCAATGCTCCAAACACAGAGGGCCGCGACACCGGACGACGCCCGGCGCGAGAGGC	719
Db	1041	TCAGGGGGGGGGCCCGACAGGGTATGAGGCCCGACACAGAGGGGCCCGCTGATCCCGCGGTAC	1100
QY	720	CCACTGCGGCGAGGGCGGAGCGGAGGACGATATCCAGAGACCGGACCGACACGCGCGCGGT	779
Db	1101	CGGGTGGGAGAGGGCCCGGCGCGCGGCGCTGGGAGACCGCGCTGTGGTGGAGAGGTGCGCGGT	1160
QY	780	GCACTGAGCGGAGACGCGCGGAGGAGCGACACCGCAAGACAGCGCGCGCAAGAGACGACCGC	839
Db	1161	G-GGAGCGTCTCGTCCGCGACGCGCCCGGCGCACGGCCGAGAGGCCCGCCCTCAGAGACGGTTC	1219
QY	840	GCAAGCCACACGCGGCTTGAAGCCCTTGCAGACGCGCGGCGACGCGCGGAGAGCGGGCGGAG	899
Db	1220	GCGGAGCGGACCGCGCCACCCCGGACGCGCGGCGCCCGCGCGCCCTTGTGGCGGTGTG	1279
QY	900	GCAACCGCAG-----GAGACCGACACGACCGAACCGGACCGGCGCAGGCGGACGAGACCGGAC	954
Db	1280	GTTCCCGGAGTCCGCGGTGCGCGGGGCGGGGTCTCGGGAGCCCGACAGGCGACCGCCCGCGGAC	1339
QY	955	G 955	
Db	1340	G 1340	

RESULT 8  
 US-09-963-333-5  
 ; Sequence 5, Application US/09963333  
 ; Patent No. 6664062  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stanton, Jr., Vincent P.  
 ; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES  
 ; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT  
 ; TITLE OF INVENTION: OF DISEASE  
 ; FILE REFERENCE: 11926-015002  
 ; CURRENT APPLICATION NUMBER: US/09/963,333  
 ; CURRENT FILING DATE: 2001-09-24  
 ; PRIOR APPLICATION NUMBER: 09/658,659  
 ; PRIOR FILING DATE: 2000-09-08  
 ; PRIOR APPLICATION NUMBER: 09/596,033

```

1  PRIOR FILING DATE: 2000-06-15
2  PRIOR APPLICATION NUMBER: 09/357,743
3  PRIOR FILING DATE: 1998-07-20
4  PRIOR APPLICATION NUMBER: 09/357,024
5  PRIOR FILING DATE: 1999-07-19
6  PRIOR APPLICATION NUMBER: 60/093,484
7  PRIOR FILING DATE: 1998-07-20
8  NUMBER OF SEQ ID NOS: 16
9  SOFTWARE: FASTSQ for Windows Version 4.0.
10 SEQ ID NO 5
11 LENGTH: 3772
12 TYPE: DNA
13 ORGANISM: Homo sapiens
14 FEATURE:
15 NAME/KEY: misc_feature
16 LOCATION: 431..441
17 OTHER INFORMATION: n = a or g
18 NAME/KEY: misc_feature
19 LOCATION: 498
20 OTHER INFORMATION: n = c or t
21 NAME/KEY: misc_feature
22 LOCATION: 579..599
23 OTHER INFORMATION: n = g or c
24 US-09-963-333-5

```

Query Match	4.4%	Score 41.8;	DB 4;	Length 3772;
Best Local Similarity	50.8%;	Pred. No. 0.049;		
Matches 153;	Conservative 0;	Mismatches 142;	Indels 6;	Gaps 2

Qy	660	TCGACGAAATGCTCAAACTACAGAGGAGCCGACAGACCGGAGGCGACCCCGGCGAGAGGC	719
Db	1041	TCAGGGGGGGGCCCCCAGAGGGGTAGAGCCCGACAGACAGAGGGCCGCGTGAACCCGAGCGTAC	1100
Qy	720	CCACTGCGCGAGGCGAGAGGCGAAGACATACAGAGACCGCAGCCACACGCGCGCCGT	779
Db	1101	CGAGTGTGGAGAGAGGCGCGCGCGCGCGCGCTGTGAGAGACGCGCTGTGAGAGGTTGCCCGT	1160
Qy	780	GCATGAGACGAGACACGAGCCGAGGAGCGACAGACCCGAGCAGCGCGCCCAAGAGACGACCG	839
Db	1161	G-GGAGCGTCTCTGCGAGGAGCGCCGAGCCACAGCGGAGAGCGCCGCGCTTACAGAGCGGTTG	1215
Qy	840	GCAAGCCACACGCGCGCTTAGGCGCTTGCGCCACAGGCTCGGACACGCGCGAAGAGCGCGGCGGA	899
Db	1220	GCGGAGCGGACCCCGGCCACCCCGAGCGCGCGCGCGCGCGCGCGCTTGTGAGCGGTGTGA	1275
Qy	900	GGAACGCGAG-----GAGACCGACACCGAACCGGACCGGCGCCAGAGGACAGACGACACCGA	954
Db	1280	GATCCGAGAGTCCGCGGTGCGCGGAGCCGAGATCCGGAGAGCCCAAGGACGCGCGCGCGGCA	1335
Qy	955	G 955	
Db	1340	G 1340	

RESULT 9 -030A-3  
US-08-406-030A-3  
Sequence 3, Application US/08406030A  
Patent No. 6270989  
GENERAL INFORMATION:  
APPLICANT: Treco, Douglas A.  
APPLICANT: Heartlein, Michael W.  
APPLICANT: Haug, Brian M.  
APPLICANT: Selden, Richard F.  
TITLE OF INVENTION: Proteiin Production and Delivery  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C  
STREET: Two Millita Drive  
City: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:



```

QY 710 CGGAGAGCCCCACTGCGCGGAGGGCGAGCGGACGATACGAGACCGGACCCAC 769
DB 687 COTCATGTCTGCGGAGGAGCGGGCTGTGGTGCCGACCATGACGCGGACCTTGGC 746
QY 770 GCGCGCCCGTGCATGAGACGAGACGCGCGGAGGAGCGGACCGGACCGGCGCAA 829
DB 747 GCATGTCCGCGAGTGGCGCGGCTACGCGATGAGCCATCGCGAGTCCCGACCACTGGA 806
QY 830 GAGGACCGCGGACCGACCGCGGCTAGGCGCTCGGCGGACGCGCGGCGGCGGAG 889
DB 807 GCGCGCGAGGCGGACCGACCGAGACGCGCTGAGTGTGATGAGCGCGGCGCAACGTGT 866
QY 890 GCGGCGCGGAGCAGCGGACCGGAGACCG 915
DB 867 CCGCGCGGCTTCCCACTCGGCGAACG 892

```

## RESULT 12

```

US-09-252-991A-11591/C
; Sequence 11591, Application US/09252991A
; Patent No. 6551795

```

## GENERAL INFORMATION:

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11591
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11591

```

```

Query Match 4.3%; Score 41.2; DB 4; Length 1485;
Best Local Similarity 50.0%; Pred. No. 0.051;
Matches 103; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```

```

QY 710 CGGAGAGCCCCACTGCGCGGAGGGCGAGCGGACGATACGAGACCGGACCCAC 769
DB 622 COTCATGTCTGCGGAGGAGCGGGCTGTGGTGCCGACCATGACGCGGACCTTGGC 563
QY 770 GCGCGCCCGTGCATGAGACGAGACGCGCGGAGGAGCGGACCGGACCGGCGCAA 829
DB 562 GCATGTCCGCGAGTGGCGCGGCTACGCGATGAGCCATCGCGAGTCCCGACCACTGGA 503
QY 830 GAGGACCGCGGACCGACCGCGGCTAGGCGCTCGGCGGACGCGCGGCGGCGGAG 889
DB 502 GCGCGCGAGGCGGACCGACCGAGACGCGCTGAGTGTGATGAGCGCGGCGCAACGTGT 443
QY 890 GCGGCGCGGAGCAGCGGACCGGAGACCG 915
DB 442 CCGCGCGGCTTCCCACTCGGCGAACG 417

```

## RESULT 13

```

US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558

```

## GENERAL INFORMATION:

```

; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rih
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/35003
; CURRENT APPLICATION NUMBER: US/09/007,005B

```

```

; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

```

```

Query Match 4.2%; Score 40.2; DB 3; Length 289;
Best Local Similarity 6.7%; Pred. No. 0.051;
Matches 17; Conservative 109; Mismatches 127; Indels 0; Gaps 0;

```

```

QY 608 GGGAGCTACAGACCTTAAAAAGTTGACCGCGGAGCGGCGGCGGAGGAGCGGCTTGCAGCA 667
DB 2 GGGGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 61
QY 668 ATGCCAACAACAGAGGCGGCGGAGCGGAGCGGACCGGCGGAGGAGGAGGAGGAGGAGG 727
DB 62 SRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 121
QY 728 GCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 787
DB 122 SRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 181
QY 788 CGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 847
DB 182 SRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 241
QY 848 CACGCGCTAGGC 860
DB 242 CRURCRURGR 254

```

## RESULT 14

```

US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344

```

## GENERAL INFORMATION:

```

; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rih
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350007
; CURRENT APPLICATION NUMBER: US/09/244,796
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 09/007,005
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; NAME/KEY: misc_feature

```

US-09-244-796-17

```
Query Match      4.2%; Score 40.2; DB 3; Length 289;
Best Local Similarity 6.7%; Pred. No. 0.051;
Matches 17; Conservative 109; Mismatches 127; Indels 0; Gaps 0
```

QY		608	G G A G C T C A C A G A C C T A A A A A A G T T A C C G C G A G G G C G C A A G A C A G C T T T G A C A G G A	667
Dd		2	G R G R G A R C C A B A R U N U R A R C R U R A R U R U T A R C B A A V U R U R A R C A B A R U R G R I R N R	61
QY		668	A T C C A A A C A C A G A G G C C G C A G C C G G A G G C G A C C C G G C G G A G G A C C C A C T G C G	727
Dd		62	S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N R	121
QY		728	G A G G G G G A G G C G A A G A G A T T C A G A G A C C C G A C C C A C A G C G C G C C C T G A T A G A	787
Dd		122	S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N R	181
QY		788	C G A G A C G G C C G A G G A C C C G A G C C G A G C G C G C C A A G C G A C C C G C G A A C C A	847
Dd		182	S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N S R N R N R	241
QY		848	C A C G C G C C T A G C C	860
Dd		242	C R U R C R U R D G R C	254

RESULT 15  
US-09-023-655-932

---

Search completed: April 26, 2004, 20:28:38

; Sequence 932, Application US/09023655  
; Patent No. 6607879

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.

APPLICANT: Jeffrey J. Seilhamer  
SUBSIDIARY: Subsam G. Sluatic

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENES

TITLE OF INVENTION:	EXPRE

NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIF  
COUNTY: HENR

COINTEL: 0  
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.22

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,655

FILING DATE: HEREWITH

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER

FILING DATE: ;

CLASSIFICATION:  
 UNCLASSIFIED

ATTORNEY/AGENT INFORMATION:  
NAME: Zeller Karen J

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION ;

TELEPHONE: (650) 855-05  
TELEFAX: (650) 845-4166

INFORMATION FOR SEO ID NO: 932

SEQUENCE CHARACTERISTICS:

LENGTH: 2376 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear

;  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: g1369836  
US-09-023-655-932

Query Match	4.2%;	Score 40;	DB 4;	Length 2376;
Best Local Similarity	47.0%;	Pred. No. 0.13;		
Matches 124;	Conservative 0;	Mismatches 140;	Indels 0;	Gaps 0;

Db	54	AGCCGACGGGGAGCGCGCCGCTCCGCTGCTGGGTCTCCGCGCTCCAGCGCGCGCGGGGCGGCCAG	11.3
QY	689	AGACCGCGCAGCGGAGCCCGCGCGCGAGAGAGCCCACTGGCGCAGGGCGAGGGGAGAGACAG	7.48
QY	749	ATTCGAGGACCGGAGCCACACGCGCGCCCGTCGTGACATGAGACGGAGACCGCGGAGGAGCGC	8.08
Db	114	ACCTTAGGCTCAGCCTTCGCGCCCGCGTCCCACTGAGAGACCGCGCGCTCCCGGCTTCGC	1.737
QY	809	AGACCCGAGACGCGCGCCAGAGAGGACCGGCGCAAGCCACACGCGCGCTTAGGCGCTTCGC	8.68
Db	174	GTCGCCGGAATGGAGCGGCGCGGGAGGACATGGACATGGAGGCCCCGAGGCGCCCGGGACCAACGCGC	2.33
QY	869	ACGGCGGCGCACGCGCGAGAGGCGGGGCGGAGGACCGGAGGAGACCGACCGACCGGACCTCGA	9.28
Db	234	GGGCGCGGCGACCGCTCTCGCGCGCTTCGCCGCCCTCCGCGGCTTGCTGCTGCGCGCGCG	2.93
QY	929	CGGCGCCGAGGGGAGGAGAGCAACC	9.52
Db	294	CGGCGCCACCAACCGGCGCGCTCC	3.17

```
Search completed: April 26, 2004, 20:28:38
Job time : 83.981 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 18:46:58 ; Search time 398.717 Seconds  
(without alignments)  
10821.600 Million cell updates/sec

Title: US-10-076-747-7  
Perfect score: 957  
Sequence: 1 ttgsgcttaactaatgcatg.....ggcagcagagccacgagct 957

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2907579 seqs, 2254313464 residues  
Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PC1\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PC1US\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10C\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	956	99.9	957	15	US-10-076-747-7 Sequence 7, Appl1
2	809.6	84.6	1460	15	US-10-076-747-8 Sequence 8, Appl1
3	55.2	5.8	349	15	US-10-074-475-147 Sequence 147, Appl
4	52.4	5.5	594	13	US-10-142-425-10 Sequence 10, Appl
5	52.4	5.5	594	15	US-10-123-155-10 Sequence 10, Appl
6	52.4	5.5	594	15	US-10-146-731-10 Sequence 10, Appl
7	52.4	5.5	594	15	US-10-140-472-10 Sequence 10, Appl
8	52.4	5.5	594	15	US-10-141-761-10 Sequence 10, Appl
9	52.4	5.5	594	15	US-10-142-885-10 Sequence 10, Appl
10	52.4	5.5	594	15	US-10-158-790-10 Sequence 10, Appl
11	52.4	5.5	594	15	US-10-137-871-10 Sequence 10, Appl
12	52.4	5.5	594	16	US-10-140-823-10 Sequence 10, Appl
13	52.4	5.5	594	16	US-10-141-756-10 Sequence 10, Appl
14	52.4	5.5	594	16	US-10-141-759-10 Sequence 10, Appl

15	52.4	5.5	594	16	US-10-140-805-10	Sequence 10, Appl
16	52.4	5.5	594	16	US-10-140-864-10	Sequence 10, Appl
17	52.4	5.5	3685	13	US-10-425-114-38362	Sequence 34362, A
18	50.2	5.2	872	15	US-10-198-846-6247	Sequence 6247, Ap
19	44.4	4.6	1185	13	US-10-425-114-2918	Sequence 2918, Ap
20	44.4	4.6	2265	16	US-10-062-674-1761	Sequence 1761, Ap
21	43.2	4.5	3600	15	US-10-198-846-13029	Sequence 13029, A
22	43.2	4.5	71292	13	US-10-087-192-1942	Sequence 1942, Ap
23	43	4.5	366	9	US-09-960-353-9109	Sequence 9109, Ap
24	43	4.5	764	9	US-09-981-353-46	Sequence 46, Appl
25	42.8	4.5	5760	13	US-10-152-886-14	Sequence 14, Appl
26	42.4	4.4	2754	13	US-10-156-761-2167	Sequence 2167, Ap
27	42.4	4.4	2919	15	US-10-282-122A-14574	Sequence 14574, A
28	42.4	4.4	8091	15	US-10-101-510-86	Sequence 86, Appl
29	42.4	4.4	8091	15	US-10-356-625-1	Sequence 1, Appl1
30	42.4	4.4	8091	16	US-10-159-563-322	Sequence 322, App
31	42.4	4.4	9025608	15	US-10-156-761-1	Sequence 1, Appl1
32	42.2	4.4	895	15	US-10-078-090-78	Sequence 78, Appl
33	41.2	4.3	1440	15	US-10-156-761-6190	Sequence 6190, Ap
34	41.2	4.3	1164	9	US-09-815-242-4078	Sequence 4078, Ap
35	41.2	4.3	1164	13	US-10-282-122A-7370	Sequence 7370, Ap
36	41.2	4.3	3066	15	US-10-121-988-152	Sequence 152, App
37	41.2	4.3	3066	15	US-10-200-562-152	Sequence 152, App
38	41.2	4.3	3066	15	US-10-237-551-152	Sequence 152, App
39	41.2	4.3	154746	10	US-09-827-688-8	Sequence 8, Appl1
40	41.2	4.3	154746	10	US-09-827-688-8	Sequence 25620, A
41	41	4.3	984	13	US-10-282-122A-25620	Sequence 119, App
42	40.6	4.2	536	17	US-10-338-110-119	Sequence 973, App
43	40.6	4.2	544	13	US-10-424-599-973	Sequence 18851, A
44	40.6	4.2	721	13	US-10-425-114-18851	Sequence 3085, Ap
45	40.6	4.2	2277	15	US-10-156-761-3085	

## ALIGNMENTS

RESULT 1  
US-10-076-747-7 Application US/40076747  
Sequence 7, Appl1  
Publication No. US20030180726A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Hu, Ping  
APPLICANT: Recipon, Hervé  
APPLICANT: Karia, Kalpana  
APPLICANT: Caffarely, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and  
FILE REFERENCE: DEX-0315  
CURRENT APPLICATION NUMBER: US/10/076,747  
CURRENT FILING DATE: 2002-02-13  
PRIOR APPLICATION NUMBER: 60/268,290  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 60/268,834  
PRIOR FILING DATE: 2001-02-15  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 7  
LENGTH: 957  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (519)..(519)  
OTHER INFORMATION: a, c, g or t  
US-10-076-747-7

Query Match 99.9% Score 956; DB 15; Length 957;  
Best Local Similarity 100.0%; Pred. No. 1.8e-293;  
Matches 957; Conservative 0; Mismatches 0; Gaps 0;

QY 1 TTGGGCTTACTAATGATGCTGAGCGGCGCCAAAGTGTGATGGATGCGTGTGCGCGC 60  
 Db 1 TTGGGCTTACTAATGATGCTGAGCGGCGCCAAAGTGTGATGGATGCGTGTGCGCGC 60  
 QY 61 CGAGGTACCGCGGTGAGCGCTATCTGAGCTTGAAGCGCTGATGAGTGTGAGTCCAC 120  
 Db 61 CGAGGTACCGCGGTGAGCGCTATCTGAGCTTGAAGCGCTGATGAGTGTGAGTCCAC 120  
 QY 121 AGAAGCTACAGCGCTGAGAGCTTATGCTGATGAAGGAGACACCGTGTGATGAGACA 180  
 Db 121 AGAAGCTACAGCGCTGAGAGCTTATGCTGATGAAGGAGACACCGTGTGATGAGACA 180  
 QY 181 GTGGTCCCTTACAGAAATGTCTATAGTGTCTTAACAGCTCAGCCCCCACTACCAATG 240  
 Db 181 GTGGTCCCTTACAGAAATGTCTATAGTGTCTTAACAGCTCAGCCCCCACTACCAATG 240  
 QY 241 GGGAGCTAGCAGCTGAGAGGATGCCAAGGAGAGGGGTCTCTGCTCACTCCACACCA 300  
 Db 241 GGGAGCTAGCAGCTGAGAGGATGCCAAGGAGAGGGGTCTCTGCTCACTCCACACCA 300  
 QY 301 AGGGCAGTCAAAAGCCCTTATCATGCGCATGTGACGTATGATAAAGCCCTACAAAT 360  
 Db 301 AGGGCAGTCAAAAGCCCTTATCATGCGCATGTGACGTATGATAAAGCCCTACAAAT 360  
 QY 361 AAGATTTCTGCACTTGGTTGAATGTCTTACATACATACAAAGACATATCTCACTA 420  
 Db 361 AAGATTTCTGCACTTGGTTGAATGTCTTACATACATACAAAGACATATCTCACTA 420  
 QY 421 CAGGAGTGCATGCATCACCCTGCGTGGGAGGAGATGCACTTGGCTCTGCGTGTCCA 480  
 Db 421 CAGGAGTGCATGCATCACCCTGCGTGGGAGGAGATGCACTTGGCTCTGCGTGTCCA 480  
 QY 481 ATGATGACTCATATGTTACACACCGGTGTGGGCGACACATATGATGATATTTGCCCG 540  
 Db 481 ATGATGACTCATATGTTACACACCGGTGTGGGCGACACATATGATGATATTTGCCCG 540  
 QY 541 GATATCAGACCATGATTTCCCGCGTGTGATGCAATATGCGGTACGCGGATCAAA 600  
 Db 541 GATATCAGACCATGATTTCCCGCGTGTGATGCAATATGCGGTACGCGGATCAAA 600  
 QY 601 TTGCGCCGGAGCTACAGACTTAAAGTTGACCGCGAGCGCGCGAAGAACAGGCTTT 660  
 Db 601 TTGCGCCGGAGCTACAGACTTAAAGTTGACCGCGAGCGCGCGAAGAACAGGCTTT 660  
 QY 661 CGAGCAATGCCAACAAGAGGCGCGCAGACCCGCGAGCGCGCGAGAGAGGCC 720  
 Db 661 CGAGCAATGCCAACAAGAGGCGCGCAGACCCGCGAGCGCGCGAGAGAGGCC 720  
 QY 721 CACTGCGGAGGCGAGAGGAGACATACAGAGACCGAGCGCAGCGCGCGCGCGT 780  
 Db 721 CACTGCGGAGGCGAGAGGAGACATACAGAGACCGAGCGCAGCGCGCGCGCGT 780  
 QY 781 CATGACGAGAGAGCGGCGAGAGAGCGCAGACCCGAGAGCGCGCAAGAGCGCG 840  
 Db 781 CATGACGAGAGAGCGGCGAGAGAGCGCAGACCCGAGAGCGCGCAAGAGCGCG 840  
 QY 841 CAAAGCACAAGCGCTTATGAGCTTATGAGCGAGCGCGCGCGAGAGAGGCGGAG 900  
 Db 841 CAAAGCACAAGCGCTTATGAGCTTATGAGCGAGCGCGCGCGAGAGAGGCGGAG 900  
 QY 901 CACCGGAGAGACCGACACAGAGACCGGCGCGAGCGCGAGAGAGCGCACCGAGCT 957  
 Db 901 CACCGGAGAGACCGACACAGAGACCGGCGCGAGCGCGAGAGAGCGCACCGAGCT 957

## RESULT 2

US-10-076-747-8  
 ; Sequence 8, Application US/10076747  
 ; Publication No. US20030180726A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Salceda, Susana  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Hu, Ping

; APPLICANT: Recipon, Herve  
 ; APPLICANT: Karia, Kalpana  
 ; APPLICANT: Cafertkey, Robert  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and  
 ; FILE REFERENCE: DEX-0315  
 ; CURRENT APPLICATION NUMBER: US/10/076,747  
 ; CURRENT FILING DATE: 2002-02-13  
 ; PRIOR APPLICATION NUMBER: 60/268,290  
 ; PRIOR FILING DATE: 2001-02-13  
 ; PRIOR APPLICATION NUMBER: 60/268,834  
 ; PRIOR FILING DATE: 2001-02-15  
 ; NUMBER OF SEQ ID NOS: 129  
 ; SOFTWARE: Patent version 3.1  
 ; SEQ ID NO 8  
 ; LENGTH: 1460  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1022)..(1022)  
 ; OTHER INFORMATION: a, c, g or t  
 US-10-076-747-8  
 Query Match: 84.6%; Score 809.6; DB 15; Length 1460;  
 Best Local Similarity 96.5%; Pred. No. 8e-247;  
 Matches 864; Conservative 0; Mismatches 19; Indels 12; Gaps 3;  
 QY 63 AGGTACGGCGTACAGAGCTATCTGAGCTGAGCGCCCTGATGAGTGTGAGTCCACAG 122  
 Db 578 AAGTACGGCGTACAGAGCTATCTGAGCTGAGCGCCCTGAGTGTGAGTCCACAG 633  
 QY 123 AAGCTACAGCTTACAGAGCTTATGCTGATGAGAGAGAGAGAGAGAGAGAGAGT 182  
 Db 634 AAGCTACAGCTTACAGAGCTTATGCTGATGAGAGAGAGAGAGAGAGAGAGAGT 685  
 QY 183 GGTGCTTACAGAGAGAGTGTCTATAGTGTCTTAAACAGCTGAGCGCCCTACCAATGGC 242  
 Db 686 GGTGCTTACAGAGAGAGTGTCTATAGTGTCTTAAACAGCTGAGCGCCCTACCAATGGC 745  
 QY 243 GAGACTAGCAGCTGAGAGGATCCCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 302  
 Db 746 GAGACTAGCAGCTGAGAGGATCCCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 805  
 QY 303 GCGAGTCAAAAGCCCTTATCATGCGCATGTGACGCTATGTAAGAGCGCTTACAAATA 362  
 Db 806 GCGAGTCAAAAGCCCTTATCATGCGCATGTGACGCTATGTAAGAGCGCTTACAAATA 865  
 QY 363 GATATCTGACCTTGTGTAATGTCTATCATATCAATTAACAAAGACATATCACTACA 422  
 Db 866 GATATCTGACCTTGTGTAATGTCTATCATATCAATTAACAAAGACATATCACTACA 925  
 QY 423 CGGAGTGCATGATCAGCGGTCCGTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGT 482  
 Db 926 CGGAGTGCATGATCAGCGGTCCGTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGT 985  
 QY 483 GATGACTCATATGTTACACACCGGTGTGGCGCGCAGACATATGAGTGTATTCGCGG 542  
 Db 986 GATGACTCATATGTTACACACCGGTGTGGCGCGCAGACATATGAGTGTATTCGCGG 1045  
 QY 543 ATCAGAGACCATGATTTCCCGCGTGTGATGCAATATGCGGTACGCGGAGATCAAT 602  
 Db 1046 ATCAGAGACCATGATTTCCCGCGTGTGATGCAATATGCGGTACGCGGAGATCAAT 1105  
 QY 603 CGCCCGGAGCTTACAGCTTAAAGTTGACCGCGCAGCGCGCGAGAGAGAGAGCTTTG 662  
 Db 1106 CGCCCGGAGCTTACAGCTTAAAGTTGACCGCGCAGCGCGCGAGAGAGAGAGCTTTG 1165  
 QY 663 ACGGAATGCCAACAAG 722  
 Db 1166 ACGGAATGCCAACAAG 1225

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filavotto, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Geritsen, Mary E

Query Match	5.5%;	Score 52.4;	DB 15;	Length 594;
Best Local Similarity	7.0%;	Pred. No. 5.2e-06;		
Matches 34;	Conservative 189;	Mismatches 265;	Indels 0;	Gaps 0;

RESULT 6  
US-10-146-731-10

Sequence 10, Application US/10146731  
Publication No. US20030129692A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Beresini, Laura  
APPLICANT: Desnoyers, Luc

Query Match	5.5%	Score 52.4	DB 15	Length 594
Best Local Similarity	7.0%	Pred. No. 5,2e-06		
Matches 34	Conservative 189	Mismatches 265	Indels 0	Gaps 0

## RESULT 7

US-10-140-472-10  
; Sequence 10, Application US/10140472  
; Publication No. US20030138088A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.



```

/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Inc
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Geo, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Collin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P933ORIC168
/ CURRENT APPLICATION NUMBER: US/10/140,472
/ PRIORITY FILING DATE: 2002-05-06
/ Prior Application removed - see File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 550
/ SEQ ID NO 10
/ LENGTH: 594
/ TYPE: PRF
/ ORGANISM: Homo Sapien
US-10-140-472-10

```

```

1      Publication No. US20030148432A1
2      GENERAL INFORMATION:
3      APPLICANT: Baker, Kevin P.
4      APPLICANT: Beresini, Maureen
5      APPLICANT: DeForge, Laura
6      APPLICANT: Desnoyers, Luc
7      APPLICANT: Filvaroff, Ellen
8      APPLICANT: Gao, Wei-Qiang
9      APPLICANT: Gerritsen, Mary E.
10     APPLICANT: Goddard, Audrey
11     APPLICANT: Godowski, Paul J.
12     APPLICANT: Gurney, Austin L.
13     APPLICANT: Sherwood, Steven
14     APPLICANT: Smith, Victoria
15     APPLICANT: Stewart, Timothy A.
16     APPLICANT: Tuma, Daniel
17     APPLICANT: Watanabe, Colin K
18     APPLICANT: Wood, William
19     APPLICANT: Zhang, Zemin
20     TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
21     FILE REFERENCE: P3330R.0198
22     CURRENT APPLICATION NUMBER: US/10/141,761
23     CURRENT FILING DATE: 2002-05-08
24     Prior Application removed - See Palm or File Wrapper
25     NUMBER OF SEQ ID NOS: 550
26     SEQ ID NO 10
27     LENGTH: 594
28     TYPE: PRT
29     ORGANISM: Homo Sapien
30     US-10-141-761-10

```

Query Match 5.5%; Score 52.4; DB 15, Length 594;  
Best Local Similarity 7.0%; Pred. No. 5,2e-06;  
Matches 34; Conservative 189; Mismatches 265; Indels 0; Gaps 0.

Oy 393 ATACATCAAGACACATPCTCACTCAGGAGTGATCGATACCGGTCGTGGGG 452  
|||||.....:::..|||.....:::..|||  
60 TTTTTTGGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT

```

; ORGANISM: Homo Sapien
US-10-141-761-10

Query Match.          5.5%; Score 52.4; DB 15; Length 594;
Best Local Similarity 7.0%; Pred. No. 5.2e-06;
Matches 34; Conservative 189; Mismatches 285; Indels 0; Gaps 0

```

Db 80 AASSVNNMTPQGNINNMIRPDKKADHDHLDVITPPLITLITSPQSSDEKKNVEXKRL 119

Qy 453 CGATGCCACTTGGCTCTGGTCGTCGTCATGATGATCATAGTTATGACACCGGTGCGG 512

Db 120 VQNDPAGISBEQCLYQIYIDELYGGLQRPSEDEKKLAEKASIGTYEDSTVAIEKAA 179

Qy 513 CGCAACNATGATAGTGAATATCGCGCGGAGTCACAGACCATGGATTTCCCCCGGCT 572

Db 180 EKEEESHAEEESNSDEEVIPLDVIDVVDDELANGEVDALINKQATTTGMDGCVIRML 239

Qy 573 GATCGAATYTCGGGTACCGGCATCAAAATTCGCCCGGAGCTACAGCTAAAAAGTTG 632

Db 240 RKQKEBAEAIKNAKLEEKAMYSGRSRRRORRRERKEKLRGKISPTAYARDSTYTP 299

Qy 633 ACCGCCAGCGGCGCAAGAACAGCTTTCGACGGAATGCCAAACACAGAGCGCGGACG 692

Db 300 YKSPSESSSEBSRSRPTPREKEKITFTSPGSGPEAAAAAASGVTTGPRAP 359

Qy 693 CGGCAGCCGACCCCGGCGGAGAGACCCCACTGCGGACGAGCGGAAGACAGCATNC 752

Db 360 PQGPGAPARNASARRSSSSSSSSASRTSSSSSSSSRSRRRGQYRRSRNHRSSAS 419

Qy 753 GAGAGACGAGGACCAACGCGCGCTCATATGAACAGGAACGAGCGGACGAGGACGAC 812

Db 420 RSMRSRSMRSRRSRNRSGRHSGGSRDGHFRSPAPRGVGPFRMRSRSHSGREY 479

Qy 813 CGGAACACCGCGGCGCAAGAGGACCGCGGACGACCAACGCGGCTAGGCTCGGCGGACG 872

Db 480 RRGGRGLRHSSSRSSWSLSPASRLTNRSHSPSPQSSRSRSRSGSPSPAPAE 539

Qy 873 CCGGCCAC 880

Db 540 KLTTPPAA 547

393	ATACATTAACAAAGACACATACTCAACTACACAGAGATCGATGATCACCGGTCCGTGGCG	452
Db	AESPVMVMFMOCDTNMIDRFVRAHLHDHIDPYPLTLTISPEQESDERKCYNERGL	119
453	CGAATGCCACTTCGCTCCGTGGCGGCGCAATGATGATGCATGTGTTAAACACCGGCTGCGG	512
Db	120 VONDFPGISBEOCLYQIYIDELGYGLCPSEDEKKLAEKASIGIYEDSTVAEYKAA	179
Qy	513 CGCACANCTYATGATGATATTTCGCCCGGGATCAACAACATGATTTCCCCCGTGGCT	572
Db	180 EKEEESAAEESNSDEVIDPIDVDYVDDELNOVADLMOATTYGMAUGDFRML	239
Qy	573 GATCGATATGGGTAAGCGGCGCATAAATTCGCCCGGAGTACAGACTAAAAAGTTG	632
Db	240 RDXKEAELIKAKKLEBKAMYSGRNRQRHREFREKLGRKISPSVARRDSPTYP	289
Qy	633 ACCGCGACGCGGCCGAAGAACAAGCTTTCAGCGAATGCCAAACACAGAGGCGCCGAC	692
Db	300 YKSPSESSSRSSRSPPTGPEREKLFTTSPFGSDPEAAAAAAGVTTGKRPAP	359
Qy	693 CGGACGACACCCCGCGGAGAGAGCCCACTGGCGAGGCGGAGCGGAAGGACAGATAC	752
Db	360 POGGPAFAPRNASARRSSSSSSSABRTSSSSSSSRSSRSPGGVYRSGHAPRS	419
Qy	753 GAGGACGGAGGACACACCGCGCCGTGCATGAGAACGAGACGGCGCGAGGAGAGGCGAC	812
Db	420 RSMRSRHSKRYSRSRSGRHRHSGGSRDBHRTSRPARRGVGPFRHSRSHSGDRY	479
Qy	813 CCGAGACGCGCGCCAAAGCGACCGGCGAAGCCACACGCGCTTGGCCCGCAG	872
Db	480 RRGGRRLRHSSSRSSRSMWLSPTSRSLTTRSHSSPSQSGRSHSRSSQSSPSPAE	539

RESULT 8  
US-10-141-761-10  
; Sequence 10, Application US/10141761

Db 540 KLTRPAAS 547







[illegible]

```

RESULT 15
US-10-140-805-10
; Sequence 10, Application US/10140805
; Publication No. US20030207417A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33081C176
CURRENT APPLICATION NUMBER: US/10/140, 805
CURRENT FILING DATE: 2002-05-07
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 10
LENGTH: 594
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-805-10

```

Query Match	5.5%;	Score 52.4;	DB 16;	Length 594;
Best Local Similarity	7.0%;	Pred. No. 5.2e-06;		
Matches 34;	Conservative 189;	Mismatches 265;	Indels 0;	Gaps 0

QY 393 ATCATPACAAACACACTACTCACTAACGAGATGATGATACACGGCGCGTGGG 452

Db 60 ABBFVMMFMQGDSTNNKIDRFVRAHLHPIPTPLPLTTTSPEQESDKCYERGL 119

QY 453 CGAATGCCACTTCGCTCGCTCGTCGTCATGATGACTATGATACACCGGATGCGG 512

Db 120 VQNDPAGISEQCLVYIYIDELVYGLQGRSEDEKKLKKKASTGYTYEDSTVAEYKA 179

QY 513 CGCAGACNTYTGATGATGATTTGCGCCGGGATACAGACCATGATTTCCCCCGGCT 572

Db 180 EKPEEBSAAEBSNSNDEDEVPIDIDVEYDVDELNQEVADLNKQATTYGMADGFYRL 239

QY 573 GATGCAATTCGCGCTACGGCGCATCAAAATGCGCCGGAGCTAACAGACTAAAAAGTTG 632

Db 240 RKQKEAEYIKAKKALEEKYMSGRSHRQREFRERLNGRTISPSYARNSPTYP 299

[illegible]

Search completed: April 26, 2004, 22:54:08  
Job time : 405.717 secs

540 KULTURA 547

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:46:02 ; Search time 2327.37 Seconds  
(without alignments)  
12279.158 Million cell updates/sec

Title: US-10-076-747-7  
Perfect score: 957  
Sequence: 1 ttgggcttactatgcacg.....ggcagagagaccacgcagct 957

Scoring table: IDENTITY NUC  
Gapox 10.0, Gapext 1.0

Searched: 27513289 seqs, 1491090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_gesba:\*  
2: em\_gesbba:\*  
3: em\_estba:\*  
4: em\_estba:\*  
5: em\_estba:\*  
6: em\_estba:\*  
7: em\_estba:\*  
8: em\_estba:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estba:\*  
16: em\_estba:\*  
17: em\_ges\_hum:\*  
18: em\_ges\_hum:\*  
19: em\_ges\_hum:\*  
20: em\_ges\_hum:\*  
21: em\_ges\_hum:\*  
22: em\_ges\_hum:\*  
23: em\_ges\_hum:\*  
24: em\_ges\_hum:\*  
25: em\_ges\_hum:\*  
26: em\_ges\_hum:\*  
27: em\_ges\_hum:\*  
28: gb\_ges1:\*  
29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.6	8.5	1018	12	BM914338 AGENCOURT
2	69.6	7.3	891	14	CD107024 AGENCOURT
3	67.4	7.0	770	14	BG686134 AGENCOURT
4	66.6	7.0	1201	13	BX356664 AGENCOURT

Result No.	Score	Query Match	Length	ID	Description
5	66.4	6.9	1201	9	AL531683 AGENCOURT
6	65.4	6.8	753	13	BU933171 AGENCOURT
7	65.4	6.8	939	29	CNS0049E AGENCOURT
8	64.8	6.8	925	29	CNS0091P AGENCOURT
9	64.6	6.8	1144	13	BX415926 AGENCOURT
10	62.6	6.5	1100	29	CNS016KD AGENCOURT
11	62.6	6.5	1213	12	BM918618 AGENCOURT
12	62.4	6.5	1101	29	CNS017SY AGENCOURT
13	61.4	6.4	844	29	CNS0052P AGENCOURT
14	61.2	6.4	545	29	BX424977 AGENCOURT
15	61.1	6.4	510	29	CNS006ON AGENCOURT
16	60.8	6.3	958	13	BU146944 AGENCOURT
17	60.2	6.3	1201	13	BX340009 AGENCOURT
18	60.1	6.3	107	13	BX457099 AGENCOURT
19	60.0	6.3	1160	10	BF129273 AGENCOURT
20	59.8	6.2	1103	13	BX403654 AGENCOURT
21	59.2	6.2	932	29	CNS0072Q AGENCOURT
22	59.1	6.2	624	13	BX404419 AGENCOURT
23	58.8	6.1	765	13	BX595659 AGENCOURT
24	58.8	6.1	1201	13	BX360624 AGENCOURT
25	58.6	6.1	932	29	CNS0072Q AGENCOURT
26	58.6	6.1	939	13	BQ706586 AGENCOURT
27	58.2	6.1	937	29	CNS006XP AGENCOURT
28	58.2	6.1	1095	9	AL551270 AGENCOURT
29	58.1	6.1	912	29	CNS006N3 AGENCOURT
30	57.6	6.0	886	13	BQ712155 AGENCOURT
31	57.6	6.0	924	13	BX442207 AGENCOURT
32	57.2	6.0	1200	13	BX415896 AGENCOURT
33	57.1	6.0	1201	29	CNS016BR AGENCOURT
34	56.8	5.9	932	29	CNS0070E AGENCOURT
35	56.6	5.9	935	29	CNS006K AGENCOURT
36	56.4	5.9	1159	29	CNS015XR AGENCOURT
37	56.2	5.9	784	29	AG161654 Pan trogl AGENCOURT
38	56.1	5.9	938	29	CNS00KX6 AGENCOURT
39	56.0	5.9	1015	28	BZ569259 pac82-164 AGENCOURT
40	55.8	5.8	1165	13	BX425080 AGENCOURT
41	55.8	5.8	1201	13	BX405071 AGENCOURT
42	55.4	5.8	1101	29	CNS00BNG AGENCOURT
43	55.2	5.8	894	29	CNS01591 AGENCOURT
44	55.2	5.8	953	13	BQ709520 AGENCOURT
45	55.2	5.8	1200	29	CNS01671 AGENCOURT

## ALIGNMENTS

RESULT 1  
BM914338 1018 bp mRNA linear EST 12-MAR-2002  
LOCUS AGENCOURT 6615480 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:5480006  
DEFINITION 5', mRNA sequence.  
ACCESSION BM914338  
VERSION BM914338.1 GI:19364717  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 1018)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strussberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Mark Watson  
CDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: L10CM2002 row: C column: 15  
High quality sequence stop: 644.

## FEATURES

source

Location/Qualifiers

1. 1018  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5480006"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_113"  
 /note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 8.5%; Score 81.6; DB 12; Length 1018;  
 Best Local Similarity 62.6%; Pred. No. 5.9e-08;  
 Matches 218; Conservative 0; Mismatches 109; Indels 21; Gaps 5;

QY 63 AGTACGGGGTCCAGAGCTATCTGAGCTGACGGCTGATGCACTGTAAGTCCACAG 122  
 DB 591 AAGTACGGGGTCCAGAGCTATCTGAGCTGACGGCTGATGCACTGTAAGTCCACAG 646  
 QY 123 AAGTACGGGGTCCAGAGCTATCTGAGCTGATGCACTGTAAGTCCAGT 182  
 DB 647 AAGTACGGGGTCCAGAGCTATCTGAGCTGATGCACTGTAAGTCCAGT 698  
 QY 183 GGTCCCTTACGAAATGTTCTATAGTTCTTAAACAGCTGACCCCTACCAATGCG 242  
 DB 699 GGGCCCCCTACGAAATGTTCTATAGTTCTTAAACAGCTGACCCCTACCAATGCG 753  
 QY 243 GAGACTACAGCTGACAGGATCCCAAGGAGAGGGGTCTCTCATCCACACACAG 302  
 DB 754 AAAATTGAAAGCTGACAGGATCCCAAGGAGAGGGGTCTCTCTCCCAACCCACAG 813  
 QY 303 GCGAGTCAAAAGCCCTTATCATGCGCATGTGACGTGATGTAAGGCGCTCAAAATTA 362  
 DB 814 GGAATTTAGGCGCCCTTCTCTCCCTTGCACCTTCAATTAAGCCCTCCCAAAAAA 869  
 QY 363 GATATTCGCACTGTTGAAATGCTTACATACATCAACAAAGCA 410  
 DB 870 AAAATTCCTCCATGTTGTCATCCGAGAGAAAAA 917

RESULT 2 891 bp mRNA linear EST 15-MAY-2003  
 LOCUS CD107024  
 DEFINITION AGNCOURT\_14020401 NIH\_MGC\_179 Homo sapiens cDNA clone  
 IMAGE:30368459 5', mRNA sequence.  
 ACCESSION CD107024  
 VERSION CD107024.1 GI:30759814  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 891)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein  
 cDNA Library Preparation: Invitrogen Corp  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: NDAM431 row: e column: 12

## FEATURES

source

High quality sequence stop: 533.  
 Location/Qualifiers

1. 891  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30368459"  
 /tissue\_type="Pituitary"  
 /lab\_host="DH10B-rom A (T1 and T5 phage resistances)"  
 /clone\_id="NIH\_MGC\_179"  
 /note="Organ: brain; Vector: pCMV-Sport6.1; Site\_1: EcoRV (destroyed); Site\_2: NotI; Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.1 kb. Library was constructed by (Invitrogen). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 7.3%; Score 69.6; DB 14; Length 891;  
 Best Local Similarity 64.1%; Pred. No. 3.1e-05;  
 Matches 159; Conservative 0; Mismatches 74; Indels 15; Gaps 3;

QY 63 AGTACGGGGTCCAGAGCTATCTGAGCTGACGGCTGATGCACTGTAAGTCCACAG 122  
 DB 594 AAGTACGGGGTCCAGAGCTATCTGAGCTGACGGCTGATGCACTGTAAGTCCACAG 649  
 QY 123 AAGTACGGGGTCCAGAGCTATCTGAGCTGATGCACTGTAAGTCCAGT 182  
 DB 650 AAGTACGGGGTCCAGAGCTATCTGAGCTGATGCACTGTAAGTCCAGT 703  
 QY 183 GGTCCCTTACGAAATGTTCTATAGTTCTTAAACAGCTGACCCCTACCAATGCG 242  
 DB 704 GGGCCCCCTACGAAATGTTCTATAGTTCTTAAACAGCTGACCCCTACCAATGCG 758  
 QY 243 GAGACTACAGCTGACAGGATCCCAAGGAGAGGGGTCTCTCATCCACACACAG 302  
 DB 759 GAAAGCTAAATCTGACAGGATCCCAAGGAGAGGGGTCTCTCTCCCAACCCCAAGC 818  
 QY 303 GCGAGTC 310  
 DB 819 ATCAAGTC 826

RESULT 3 770 bp mRNA linear EST 01-MAY-2001  
 LOCUS BG686194/c  
 DEFINITION 602638481F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4766108 5',  
 mRNA sequence.  
 ACCESSION BG686194  
 VERSION BG686194.1 GI:13917591  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 770)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LILCM1626 row: a column: 21  
 High quality sequence stop: 307.  
 Location/Qualifiers

## FEATURES

source

1. 770  
 /organism="Homo sapiens"

```

1. 1201
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /dd_xref="taxon:9606"
   /clone="CC001015YB03"
   /tissue_type="PLACENTA COT 25-NORMALIZED"
   /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
   /note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM0081701"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and

```









/lab host="DH10B"  
 /clone.lib="NIH\_MGC\_120"  
 /note="Organ: pooled pancreas and spleen; Vector:  
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 kb, insert size range 1-2.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 6.5%; Score 62.6; DB 12; Length 1213;  
 Best Local Similarity 57.1%; Pred. No. 0.0014;  
 Matches 214; Conservative 0; Mismatches 149; Indels 12; Gaps 5;

QY 70 CGGTACAGAGCTATCTAGAGCTTACAGGCTGATGAGAGTCCACAGAGCTAC 129  
 DB 623 CGGCGACAGCTATCTAGAGCTTACAGGCTGATGAGAGTCCACAGAGCTTA 680  
 QY 130 AGCTGACAGAGCTTATGCTGATGAGAGACACCTGATGAGAGAGTGT-CCC 188  
 DB 661 CAGCTGCGAGAGCTGACGATG--GAGGGAGACCTGAGAGAGAGAGGCGCC 737  
 QY 189 CTACAGAAATGTTCTATAGTTCTCTAACAAGCTCAGCCCTCACTACCAATGGCGAGCT 248  
 DB 738 CTACAGAAATGTTCTATAGTTCTCTAACAAGCTCAGCCCTCACTACCAATGGCGAGCT 795  
 QY 249 AGCAGCTGCGAGAGTCCCAAGAGAGGG---TCTCTCATCAACACCAAGAGG 304  
 DB 796 AAAAATGCTCCAGAGATCCCAAGAGAGGGATTTCTCTCTCCCAACCCCTCCAGG 855  
 QY 305 CGAGCAAGAGCCCTATCATCGCGAGATGTCAGCTCATGTAAAGCCCTCAAAATAGA 364  
 DB 856 GCAATGCAAGAGCCCTTCTCTCTGAAATCCAAATTAACCCCTTCTCATTAATA 915  
 QY 365 TATCTGCACTGTTGAATGCTCTATCATCATACAGACACATATCTCACTACAG 424  
 DB 916 ATTCTCTCAATGGGCGCAACCCGCAAGTACAAATAAAGCGCTGATTAACCAACAA 975  
 QY 425 GAGTGCATGATCAG 439  
 DB 976 TCGCGGTTGGAAC 990

RESULT 12  
 CNS017SY 1101 bp DNA linear GSS 26-JUN-1999  
 LOCUS  
 DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC  
 BACN37L08 of DrosBAC library from Drosophila melanogaster (fruit  
 fly), genomic survey sequence.  
 AL108460  
 AL108460.1 GI:5628764  
 GSS.  
 ORGANISM Drosophila melanogaster (fruit fly)  
 SOURCE Drosophila melanogaster  
 KEYWORDS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 1101)  
 REFERENCE  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-JUN-1999) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 COMMENT Determination of this BAC-end sequence was carried out as part of a  
 collaboration with the European Drosophila Genome Project (EDGP) -  
 http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC  
 library (Dros BAC) was made by Alain Billard at CEPR (Centre  
 d'Etude du Polymorphisme Humain) with funding provided by a MGC  
 project grant. The DNA was prepared from embryos by Alain Bucheton  
 and Genevieve Payan. It has been constructed in the vector

FEATURES  
 source  
 location/Qualifiers  
 1..1101  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACN37L08"  
 /clone.lib="DrosBAC"  
 /plasmid="pBelobAct11"  
 /note="end : Sp6"

## ORIGIN

Query Match 6.5%; Score 62.4; DB 29; Length 1101;  
 Best Local Similarity 17.1%; Pred. No. 0.0015;  
 Matches 64; Conservative 178; Mismatches 129; Indels 3; Gaps 1;

QY 583 GCGTACCGCGCATCAATTGCCCCGAGAGCTACAGACCTTAAAGTTAGCCGCGCAGC 642  
 DB 667 GSGGSSCCGKAKGVGRVCCAGGASCAAMAADCGCCAKMACSSSSSSSSGSC 726  
 QY 643 GGGCGAAGACAGGCTTTCAGAGCAATGCCAAACACAGAGGCGCGACCGCGAGGGA 702  
 DB 727 AATSSASRGMVSSCAGSGSAGSAGSAGSGSGGSGSCASGVCAGACSA 786  
 QY 703 CCCCGGCGAGAGAGCCCACTGCGGAGGCGGAGCGGAGACAGATACAGAGAGCGGA 762  
 DB 787 SCSSASMGVSSGSSCSAGSCGVSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGS 846  
 QY 763 GCCACAGCGCGCCCGCTGATGAGAGAGAGAGCGCCGAGAGAGCGACAGCCGAGCGAG 822  
 DB 847 GAVSSSCRSVSVSAASVSSSVSSSSSVSAASVSSSSASASVAAAAVVA 906  
 QY 823 GCGCAAGAGAGCGCGAGAGCGGAGAGAGCGGCTGAGCCCTGCGGAGCGGCGGCGAGC 882  
 DB 907 VESVVA---SVSSSSSSSSSSSVSVSVSVSVSVSVSVSVSVSVSVSVSVSVSVSV 963  
 QY 883 GCGAGAGCGGCGGAGAGAGCGGAGAGAGCGGAGAGAGCGGAGAGCGGAGAGAG 942  
 DB 964 AVMSAVVSSSSASSSSVSVSVVAAAAAASSSSSASASVAVVSVSSSSSS 1023  
 QY 943 CAGAGCCAGCGAGC 956  
 DB 1024 ASSSSSSSSASSVS 1037

RESULT 13  
 CNS0052P 844 bp DNA linear GSS 03-JUN-1999  
 LOCUS  
 DEFINITION Drosophila melanogaster genome survey sequence Tm3 end of BAC #  
 BACR1P16 of RPCI-98 library from Drosophila melanogaster (fruit  
 fly), genomic survey sequence.  
 AL056652  
 AL056652.1 GI:4932342  
 GSS.  
 ORGANISM Drosophila melanogaster (fruit fly)  
 SOURCE Drosophila melanogaster  
 KEYWORDS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 844)  
 REFERENCE  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 COMMENT Determination of this BAC-end sequence was carried out as part of a  
 collaboration with the Berkeley Drosophila Genome Project (BDGP).  
 The BDGP is constructing a physical map of the Drosophila  
 melanogaster genome using these BACs. For further information  
 please see http://www.fruitfly.org The BDGP Drosophila  
 melanogaster BAC library was prepared by Kazuhiro Osoegawa and  
 Aaron Mammosser in Peter de Jong's laboratory in the Department of  
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo.

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA005ZH04"
/tissue_type="PLACENTA"
```

```
/db_xref="taxon:7227"
/clone="BACR14J21"
/clone_1b="RPCI-98"
/note="end : T7"
```

ORIGIN



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 14:33:42 ; Search time 5602.61 Seconds

(without alignments)  
11294.891 Million cell updates/sec

Title: US-10-076-747-8

Perfect score: 1460

Sequence: 1 ggcctgggctctgtatccc.....ggcagcagagccacgcagct 1460

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_da:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vt:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hu:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vt:\*  
30: em\_hg\_hum:\*  
31: em\_hg\_inv:\*  
32: em\_hg\_other:\*  
33: em\_hg\_mus:\*  
34: em\_hg\_pln:\*  
35: em\_hg\_rod:\*  
36: em\_hg\_mam:\*  
37: em\_hg\_vrt:\*  
38: em\_sy:\*  
39: em\_hg\_hum:\*  
40: em\_hg\_mus:\*  
41: em\_hg\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	642.2	44.0	755	9	HSIGVL033	X57823 Human reart
2	634.2	43.4	895	9	BC033102	BC033102 Homo sapi
3	626	42.9	922	9	BC062711	BC062711 Homo sapi
4	607	41.6	891	6	AK135365	AK135365 Sequence
5	606.8	41.6	811	9	AB064153	AB064153 Homo sapi
6	600.4	41.1	789	9	AB064207	AB064207 Homo sapi
7	597.4	40.9	877	9	HUMIGHPEPAL	M87790 Human (hybr
8	591.4	40.5	1230	9	AK130519	AK130519 Homo sapi
9	588.4	40.3	894	9	BC018749	BC018749 Homo sapi
10	584.6	40.0	756	9	HSIGVL034	X57824 Human reart
11	584.4	40.0	810	9	AB064158	AB064158 Homo sapi
12	581.4	39.8	915	9	BC015833	BC015833 Homo sapi
13	570.2	39.1	783	9	HSIGVL001	X57801 Human reart
14	570	39.0	806	9	AB064155	AB064155 Homo sapi
15	570	39.0	810	9	AB064163	AB064163 Homo sapi
16	568.4	38.9	807	9	AB064154	AB064154 Homo sapi
17	568.4	38.9	808	9	AB064164	AB064164 Homo sapi
18	568.4	38.9	916	9	BC030983	BC030983 Homo sapi
19	568.2	38.9	812	9	AB064160	AB064160 Homo sapi
20	566.8	38.8	808	9	AB064152	AB064152 Homo sapi
21	565.4	38.7	729	9	AY172960	AY172960 Homo sapi
22	563.6	38.5	809	9	AB064156	AB064156 Homo sapi
23	562	38.5	805	9	AB064161	AB064161 Homo sapi
24	560.2	38.4	810	9	AB064159	AB064159 Homo sapi
25	559	38.3	885	9	AK129585	AK129585 Homo sapi
26	557	38.2	826	9	AB064215	AB064215 Homo sapi
27	550.6	37.7	807	9	AB064162	AB064162 Homo sapi
28	544.8	37.3	808	9	AB064157	AB064157 Homo sapi
29	543	37.2	887	9	AK130461	AK130461 Homo sapi
30	537.6	36.8	663	9	HS007991	U07991 Human lymph
31	531.6	36.4	885	9	AK129603	AK129603 Homo sapi
32	520.4	35.6	915	9	BC012876	BC012876 Homo sapi
33	518.4	35.5	942	6	AX780304	AX780304 Sequence
34	511.8	35.1	2112	6	AX523520	AX523520 Sequence
35	510.6	35.0	666	9	HS007992	U07992 Human myelo
36	500.2	34.3	903	9	BC020233	BC020233 Homo sapi
37	499.8	34.2	919	9	BC022098	BC022098 Homo sapi
38	498.4	34.1	935	6	AR135362	AR135362 Sequence
39	498.2	34.1	819	9	HSIGVL029	X57819 Human reart
40	497.2	34.1	605	6	AX379222	AX379222 Sequence
41	489.6	33.5	9472	6	AX287808	AX287808 Sequence
42	485.4	33.2	652	9	HUMIGHVCK	M98325 Homo sapien
43	485.2	33.2	816	9	AB064145	AB064145 Homo sapi
44	482	33.0	870	6	BD176841	BD176841 A method
45	482	33.0	870	9	HSIGLV	X14583 Human mRNA

#### ALIGNMENTS

RESULT 1  
LOCUS HSIGVL033 755 bp mRNA linear PRI 04-NOV-1994  
DEFINITION Human rearranged immunoglobulin lambda light chain mRNA.  
ACCESSION X57823  
VERSION X57823.1 GI.33745  
KEYWORDS Ig lambda light chain; immunoglobulin.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 344 to 390)  
Combratio, G. and Klobbeck, H. G.  
V lambda and J lambda-C lambda gene segments of the human  
immunoglobulin lambda light chain locus are separated by 14 kb and

**source**

Die

Db

ORGANISM

## AUTHORS



## REMARK

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Louis Staudt  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Buterfield,  
 Susanna Chan, Readman Chin, Chris Fjell, Erin Garland, Ran Guin,  
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
 Sen Lee, Victor Ling, Carrie Matheson, Candice McLeavy, Steven  
 Nees, Pawan Parдох, Anna-Lies Prabhu, Parvaneh Saeedi, Jacqueline  
 Schein, Diane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott,  
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>  
 Series: TRL Plate: 43 Row: 9 Column: 22  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Genomescan gene  
 prediction, similarity but not identity to protein.

## FEATURES

## source

Location/Qualifiers  
 1..895  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:45681 IMAGE:4851128"  
 /tissue\_type="Primary B-Cells from Tonsils"  
 /clone\_id="N1H MGC\_48"  
 /lab\_host="DH105-R"  
 /note="vector: pOTB7"  
 31..738  
 /codon\_start=1  
 /product="Unknown (protein for MGC:45681)"  
 /protein\_id="AAH3102.1"  
 /db\_xref="GI:21619848"  
 /translation="MAAAILLLTLTNGTSMASQALTOPPSAGSGSGSTTSCGT  
 SSDVGYNYSWYQHPKAPKLTIVYNNRPSGVPRPSGSSGNTATVGLAE  
 DEADYCSYAGSNVYFGTGTAVLGPAPNPTVTLFPSSSELDQANATVCLIS  
 DFYGAIVAMKADGSPKAGVETTKPSKSNKYASVLSLTPROMSHRSYSCQV  
 THESGTEKTVAPTECS"

## CDS

misc\_feature  
 139..417  
 /note="IGV; Region: Immunoglobulin domain variable region  
 (v) subfamily"  
 /db\_xref="CDD:c00099"  
 436..717  
 /note="IGC; Region: Immunoglobulin domain constant region  
 subfamily"  
 /db\_xref="CDD:c00098"

## misc\_feature

## ORIGIN

Query Match 43.4%; Score 634.2; DB 9; Length 895;  
 Best Local Similarity 95.8%; Pred. No. 4,7e-134;  
 Matches 662; Conservative 0; Mismatches 28; Indels 1; Gaps 1;

1 GAGCGGGGCTGCTATCTCCACCCCTCCTCACTCAGGACAGAGGCTCGGGGCCAGTC 60  
 33 GAGCGGGGCTGCTCTCCCTCCACCCCTCCTCACTCAGGACAGAGGCTCGGGGCCAGTC 92  
 61 TGCCCTGACAGTGTGCTGCTCGGTCTCTCTGAGACAGTGCATCCATCTCTG 120  
 93 TGCCCTGACAGTGTGCTGCTCGGTCTCTCTGAGACAGTGCATCCATCTCTG 152  
 121 CACTGGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 180  
 153 CACTGGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 212  
 181 AGGCAAAAGCCCCAAATCATATTATGAGTCAAGTAAATGAGGCTCAGAGGGTTTCTAA 240

DB 213 AGGCAAAAGCCCCAAATCATATTATGAGTCAAGTAAATGAGGCTCAGAGGGTTTCTCA 272  
 QY 241 TCGCTTCTGCTGCTCAAGTGTGGAACAACGAGCTCTCTGACCATCTCTGAGCTCAAGC 300  
 DB 273 TCGCTTCTGCTGCTCAAGTGTGGAACAACGAGCTCTCTGACCATCTCTGAGCTCAAGC 332  
 QY 301 TGAGGACGAGGCTGATTTATTAATGCTGCTCATATCAAGGATCTTCTCATGCTCTGG 360  
 DB 333 TGAGGATGAGGCTGATTTATTAATGCTGCTCATATCAAGGATCTTCTCATGCTCTGG 392  
 QY 361 AACTGGACCAAGTCAACGCTCTGATGACGACCCCAAGCCCAACCCACTGTCACTCTGTT 420  
 DB 393 AACTGGACCAAGTCAACGCTCTGATGACGACCCCAAGCCCAACCCACTGTCACTCTGTT 452  
 QY 421 CCGCGCTCTCTTGAAGAGCTCCAGCCCAAGGACCACTAGTGTCTGATAGTGA 480  
 DB 453 CCGCGCTCTCTTGAAGAGCTCCAGCCCAAGGACCACTAGTGTCTGATAGTGA 512  
 QY 481 CTTCTACCGGAGCTGTGACAGTGTGCTGGAAGGCAATGAGCAGCCCGCTCAAGCGAG 540  
 DB 513 CTTCTACCGGAGCTGTGACAGTGTGCTGGAAGGCAATGAGCAGCCCGCTCAAGCGAG 572  
 QY 541 AGTGAAGACCAACCAACCTCTCAACAGAGCAACCAAGTACGCGGCAAGCACTTACT 600  
 DB 573 AGTGAAGACCAACCAACCTCTCAACAGAGCAACCAAGTACGCGGCAAGCACTTACT 632  
 QY 601 GAGCCTGAGCCCGGAGCAAGTGAAGTCCCAAGAGCTCAAGCTGCGCAGGACGATGA 660  
 DB 633 GAGCCTGAGCCCGGAGCAAGTGAAGTCCCAAGAGCTCAAGCTGCGCAGGACGATGA 692  
 QY 661 AGGAGCAACCGTGATGAAGACAGTGTGCC 691  
 DB 693 AGGAGCAACCGTGATGAAGACAGTGTGCC 722

## RESULT 3

LOCUS BC062711 922 bp mRNA linear PRI 26-NOV-2003  
 DEFINITION Homo sapiens CDNA clone MGC:72002 IMAGE:30350364, complete cds.  
 ACCESSION BC062711  
 VERSION BC062711.1 GI:38541896  
 KEYWORDS MGC.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 922)  
 AUTHORS Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner J., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhac N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,  
 Scheetz T.E., Brownstein M.J., Udell T.B., Toschinsky S.,  
 Carrinci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J.,  
 Abramson R.D., Wulian J.S.V., Bosak S.A., McSwain P.J.,  
 McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,  
 Worley K.C., Hale S., Garcia A.M., Gay L.J., Bully S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahy J., Helton E., Kettelman M., Madan A., Young A.C., Shevchenko Y.,  
 Sanchez A., Whiting M., Madan A., Touchman J.W., Green E.D.,  
 Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D.,  
 Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 Butlerfield Y.S., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 Scheraga A., Schein J.E., Jones S.J. and Marra M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 MEDLINE 22388257  
 PUBMED 12477932  
 REFERENCE 2 (bases 1 to 922)  
 AUTHORS Strusberg R.

# TITLE JOURNAL

Direct Submission  
Submitted (25-NOV-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@cgsc.bc.ca](mailto:info@cgsc.bc.ca)

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Seedi, Jacqueline  
Shein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott,  
Michael Thorne, Miranada Tsai, Natalja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>  
Series: IRAL Plate: 51 Row: b Column: 14  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Genomescan gene  
prediction, similarity but not identity to protein.

FEATURES  
source  
location/Qualifiers

1..922  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:72002 IMAGE:30350364"  
/tissue\_type="Glandular pool- thyroid, parathyroid,  
adrenal cortex, pineal gland, submandibular gland."  
/clone\_11b="NIH MGC\_184"  
/lab\_host="DH10B"  
/note="Vector: pDNR-LIB"  
47..757  
/codon\_start=1  
/product="Unknown (protein for MGC:72002)"  
/protein\_id="AAH6711.1"  
/db\_xref="GI:38541897"  
/translation="MAALLLLTLTQTGSGMAOSALTOPASVSGSPQSITTCGT  
SSDVGYDFVSMYQONPGKAPKLIIVDSNRPSSGSKSGNTASLITGLGLE  
DEADYSGSYSTTPPEFVSGSKVTIVGGKAPVTYLPPESEELQANKATVLCI  
SDRYGATVAMKADGSPVKAQVETKSKSNKKYAASVLSLTPEQWKSHTSYSCQ  
VTHESTVEKTVAPTECS"

misc\_feature

131..436  
/note="IGV; Region: Immunoglobulin domain variable region  
(v) subfamily"  
/db\_xref="CDD:cd00099"  
455..736  
/note="IGC; Region: Immunoglobulin domain constant region  
subfamily"  
/db\_xref="CDD:cd00098"

ORIGIN

Query Match 42.9%; Score 626; DB 9; Length 922;  
Best Local Similarity 95.8%; Pred. No. 3.5e-133;  
Matches 665; Conservative 0; Mismatches 25; Indels 4; Gaps 2;  
QY 1 GGCCTGGGCTGTCTATTCTCAACCTCTCTCACTAGGCGCACAGGGTCTGGGCCCAATC 60  
DB 49 GGCCTGGGCTGTCTATTCTCAACCTCTCTCACTAGGCGCACAGGGTCTGGGCCCAATC 108  
QY 61 TGCCTGACTAGTGTGCTCGTGTCTGGGCTCTGAGACAGTGCATCACTCTCTG 120  
DB 109 TGCCTGACTAGTGTGCTCGTGTCTGGGCTCTGAGACAGTGCATCACTCTCTG 168  
QY 121 CACTGAAACGAGCTCAGCTTGCTGTATTAATACTATGTCTCTGTGATCAACAGCAACC 180

DB 169 CACTGAAACGAGCTCAGTGTGTGATATATCTTGTCTCTGTATCAACAAACCC 228  
QY 181 AGGCAAGCCCCCAACTCATATTATAGGTAGTAATCGGCCCTCAAGGGTTCTTA 240  
DB 229 AGGCAAGCCCCCAACTCATATTATAGGTAGTAATCGGCCCTCAAGGGTTCTCA 288  
QY 241 TGCCTTCTGTGCTGCAATCTGCGAACACGGGCTTCCCTGACCATCTTGGGCTCAAGC 300  
DB 289 TGCCTTCTGTGCTGCAATCTGCGAACACGGGCTTCCCTGACCATCTTGGGCTCAAGC 348  
QY 301 TGAAGCAGAGGCTGATTATTAAGTCTGCTCATATACAGAAAGTACT--TCTCATGTCTT 357  
DB 349 TGAAGCAGAGGCTGATTATTAAGTCTGCTCATATACAGAAAGTACT--TCTCATGTCTT 408  
QY 358 CGGAATCTGGACCAAGGTACCGTCTTAGGTACGCCCAAGCCCAACCCCATGTCACTT 417  
DB 409 CGGATCTGGACCAAGGTACCGTCTTAGGTACGCCCAAGCCCAACCCCATGTCACTT 468  
QY 418 GTTCCCGCCCTCCTTGAGAGCTCCAGGCAAGCCCAACAGCCCACTAGTGTCTATGAG 477  
DB 469 GTTCCCGCCCTCCTTGAGAGCTCCAGGCAAGCCCAACAGCCCACTAGTGTCTATGAG 528  
QY 478 TGACTTCAACCCGAGAGCTGTGACAGTGGCTGAAAGGCAATGGCAAGCCCGTCAAGC 537  
DB 529 TGACTTCAACCCGAGAGCTGTGACAGTGGCTGAAAGGCAATGGCAAGCCCGTCAAGC 588  
QY 538 GGAAGTGGAGACCAACCAACCTCCCAACAGAGCAACCAATGACGGGCGAGAGCTA 597  
DB 589 GGAAGTGGAGACCAACCAACCTCCCAACAGAGCAACCAATGACGGGCGAGAGCTA 648  
QY 598 CTTGACCTGAGACCCCGAGAGCTGAGAGTCCCAAGAGCTACAGTGCAGGTACGCA 657  
DB 649 CTTGACCTGAGACCCCGAGAGCTGAGAGTCCCAAGAGCTACAGTGCAGGTACGCA 708  
QY 658 TGAAGGAGACACCGTGTGATGAAGACAGTGTCTCC 691  
DB 709 TGAAGGAGACACCGTGTGATGAAGACAGTGTCTCC 741

RESULT 4

AR135365 891 bp DNA linear PAT 16-JUN-2001  
LOCUS AR135365  
DEFINITION Sequence 23 from patent US 6135941.  
ACCESSION AR135365  
VERSION AR135365.1 GI:14476037  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE  
1 (bases 1 to 891)  
Hillman,J.L., Lal,P., Tang,Y.Tom., Yue,H., Au-Young,J.,  
Corley,N.C., Guegler,K.J. and Baughn,M.R.  
TITLE Human immune system associated molecules  
JOURNAL Patent: US 6135941-A 23 24-OCT-2000;  
FEATURES  
source  
1..891  
/organism="Unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 41.6%; Score 607; DB 6; Length 891;  
Best Local Similarity 93.3%; Pred. No. 7.6e-128;  
Matches 645; Conservative 0; Mismatches 45; Indels 1; Gaps 1;  
QY 1 GGCCTGGGCTGTCTATTCTCAACCTCTCTCACTAGGCGCACAGGGTCTGGGCCCAATC 60  
DB 36 GGCCTGGGCTGTCTATTCTCAACCTCTCTCACTAGGCGCACAGGGTCTGGGCCCAATC 95  
QY 61 TGCCTGACTAGTGTGCTCGTGTCTGGGCTCTGAGACAGTGCATCACTCTCTG 120  
DB 96 TGCCTGACTAGTGTGCTCGTGTCTGGGCTCTGAGACAGTGCATCACTCTCTG 155

Qy	12	CAC	TGGAACAGAGACGACGTTGGGGTTAATACATATGTCCTCCGGAAACCAACAGACACC	180
Db	156	CA	TTGAAACAGAGATGACGTGTGGTTAAACATAGTCTCCGTGACCAACAGTCCC	215
Qy	181	AGG	CAAAAGCCCCCAAATCATTCATTTATGAGTCAAGTAATCGGCCCTCAGGGATTTC	240
Db	216	AGG	CAAGGCCGCCCAAACTCATGATTTTATGAGGTCAATATCGGCCCTCAGGGATTTC	275
Qy	241	TGC	CTTCTCTGGGTCCTCAAGTCTGGGCAACAGGGCTCCCTGACATCTCTGGGCTCAAGC	300
Db	276	TCC	GTTCCTGTGGTCCAACTCTGGGCAACAGGGCTCCCTGACATCTCTGGGCTCAAGC	335
Qy	301	TGAG	ACGAGGCTGATTATTAATGCTGCATATACAAAGAAAGTACTTCATGTCCTGG	360
Db	336	TGAG	ACGAGGCTGATTATTAATGCTGCATATATGAGCAACAACTGTGTTATTCG	395
Qy	361	AAC	TGAGACCAAGGTCAACCGGCTTAAGTCAACCCCAAGGCCAACCCCATGTCACTGT	420
Db	396	CGA	GGAGGACCAAGCTGACCGGCTTAAGTCAACCCCAAGGCCAACCCCATGTCTGT	455
Qy	421	CCC	GCCCTCTCTCTGAGGAGCTCCAAAGCCAACAAGGCCACACTAGTGTGTGATCA	480
Db	456	CCC	GCCCTCTCTCTGAGGAGCTCCAAAGCCAACAAGGCCACACTAGTGTGTGATCA	515
Qy	481	CTT	CTACCCGGGAGCTGTGACAGTGTGGCTTGAAGGACAGATGACAGCCCTCAAGCCGG	540
Db	516	CTT	CTACCCGGGAGCGGTGACAGTGTGGCTTGAAGGACAGATGACAGCCCTCAAGCCGG	575
Qy	541	AGT	GAGACCAACCAAACTCTCAAAACAGACAAACAAGTACCGGGCCAGACGTA	600
Db	576	AGT	GAGACCAACCAACCTCTCAAAACAGACAAACAAGTACCGGGCCAGACGTA	635
Qy	601	GAC	CTGACGCCCGACAGTGAAGTCCCAAGAGCTAACAGTCCCAAGGTCAAGCATGA	660
Db	636	GAC	CTGACGCCCTGACAGTGAAGTCCCAAGAGCTAACAGTCCCAAGGTCAAGCATGA	695
Qy	661	AGG	AGACACCGGTGATGAAGACAGTGTCCC	691
Db	696	AGG	AGACACCGGTGA-GAAGACAGTGTCCC	725

RESULT 5	
AB064153	
LOCUS	811 bp mRNA linear prt 02-JUL-2004
DEFINITION	Homo sapiens IGL mRNA for immunoglobulin lambda light chain V <sub>L</sub> J
ACCESSION	AB064153
VERSION	AB064153.1 GI:21669512
KEYWORDS	
SOURCE	
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Ekayayca; Meraza; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 Akahori, Y., Iba, Y., Morino, K., Shinohara, M., Hirono, Y., Yakita, M., Suzuki, K., Torii, H., Ukai, Y., Honda, T., Katsumi, H., Okada, J., Mura, K. and Kurosawa, Y.
TITLE	Construction and characterization of antibody libraries: isolation of therapeutic human antibodies and application to functional genomics
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 811)
AUTHORS	Kurosawa, Y.
TITLE	Direct Submssion
JOURNAL	Submitted (25-JUN-2001) Yoshikazu Kurosawa, Institute for Comprehensive Medical Science, Fujita Health University; Kutsukake-cho, Toyosake 470-1192, Japan (E-mail: kurosawa@fujita-hu.ac.jp, Tel:81-562-93-9387)
COMMENT	Please visit our web site
FEATURES	URL: <a href="http://www.fujita-hu.ac.jp/immunity/">http://www.fujita-hu.ac.jp/immunity/</a> .
SOURCE	Location/Qualifiers
	1..811
	/organism="Homo sapiens"

```

gene
CDS
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="U13"
    /clone_id="AIMS4"
    /note="mixture of tissues:tonsils, umbilical cords,
    peripheral blood and bone marrow"
    1..811
    /gene="IGL"
    1..>811
    /gene="IGL"
    /codon_start=1
    /product="immunoglobulin lambda light chain VLJ region"
    /protein_id="BAC01781.1"
    /db_xref="GI:21669513"
    /translation="MYTLPTAAAGGILLAAOPMAQSAVLTQPAVSQSPQSTTSG
    TGRSSDVGATNVIVSWIQHPKAPKMLITVYSKRPSSVSNRPSSGSKGNRTASLTISG
    QADEADAVYCSSTYSSTTVFGTGTQVTVLQPEKAPNTVLFPSSSELDANKATLVQ
    LIDFPYGAATVAMKADGSPKAGVETIKPSKQSNNTKPAASSTYSLTPEDQWKSRSYS
    CQWHEGSTVEKTVAPTECSAROSTPFVCEYQCGSSDLPQFPVNAAGGSGGG"
    1..66
    /gene="IGL"
    /note="pa1b signal peptide"

```

Query Match	Similarity	41.6%	Score 606.8	DB 9	Length 811
Best Local	Similarity	97.2%	Pred. No. 8,4e-128		
Matches	628	Conservative	0	Mismatches	17
				Indels	1
				Gaps	1
Qy	46	GTCTGGGCCCCAGTCTGCGCCCTGACTGATGCTGCTCCGTGTCTGGGTCTCTGACATGC	105		
Db	57	GGCCATGGCCCAAGCTCTCCCTGACTCAGCTGCTCCGTGTCTGGGTCTCTGACATGC	116		
Qy	106	GATCACCATCTCCCTGCACTGGAACACAGCAGCTGGTGGTATATACATATGATCTCTG	165		
Db	117	GATACCACTCTCTGCACTGGAACACAGCAGCTGGTGGTATATACATATGATCTCTG	176		
Qy	166	GTACCAACAGACCCCAAGCAAAAGCCCCCAACTCATATTATATAGGTCAAGTATCGGCC	225		
Db	177	GTACCAACAGACCCCAAGCAAAAGCCCCCAACTCATATTATATAGGTCAAGTATCGGCC	236		
Qy	226	CTGAGGGGGTTTCTTAATGCTCTCTCTGCTGCTGCAAGTCTGGCAACAGCGCTCCCTGACAT	285		
Db	237	CTGAGGGGGTTTCTTAATGCTCTCTCTGCTGCTGCAAGTCTGGCAACAGCGCTCCCTGACAT	296		
Qy	286	CTCTGGGCTCCAGGCTGAGGACAGAGGGTGAATTATCTGCTGCTCATATACAAAGATAC	345		
Db	297	CTCTGGGCTCCAGGCTGAGGACAGAGGGTGAATTATCTGCTGCTCATATACAAAGATAC	356		
Qy	346	TTCTCATGTCTTCGGAACTGGGACCAAAGTCAACGTCTTAGTATAGCCCAAGGCCAACCC	405		
Db	357	CACCTAATGTCTTCGGAACTGGGACCAAAGTCAACGTCTTAGTATAGCCCAAGGCCAACCC	416		
Qy	406	CACGTCACTCTGTTCGCCGCTCTCTCTGAGGAGCTCCAGGCCAACAGGCCACACTAGT	465		
Db	417	CACGTCACTCTGTTCGCCGCTCTCTCTGAGGAGCTCCAGGCCAACAGGCCACACTAGT	476		
Qy	466	GTGTCTGATGATGACTTCTTACCCGGAGCTGTGACAGTGGCTCTGGAAGCAGATGGCAG	525		
Db	477	GTGTCTGATGATGACTTCTTACCCGGAGCTGTGACAGTGGCTCTGGAAGCAGATGGCAG	536		
Qy	526	CCCCGTAAAGCGGGAATGGAGACACCAAAACCTCCAAACAGGACCAACAAAGTACGC	585		
Db	537	CCCCGTAAAGCGGGAATGGAGACACCAAAACCTCCAAACAGGACCAACAAAGTACGC	596		
Qy	586	GGCCAGCAGCTACTGAGCTTGAAGCCCGAGCAGATGGAAAGTCCCAAGAACTTACAGCTG	645		
Db	597	GGCCAGCAGCTACTGAGCTTGAAGCCCGAGCAGATGGAAAGTCCCAAGAACTTACAGCTG	656		
Qy	646	CCAGGTACCGCATGAAGGAGACACCGTGTGATGAAGAAGTGTCTCC	691		
Db	657	CCAGGTACCGCATGAAGGAGACACCGTGTGATGAAGAAGTGTCTCC	701		

[illegible]

D	b		237	CTCAGGGGTTTCATATCGCTTCTTGSGCTCCAACTGTGGCAACAGGCTCTCCGACCAT	296
Q	y		286	CTCTGGGCTCCAGGCTGAGAACGAGGCTGAATTATTACTGCTCATATACAAGAATA	345
D	b		297	CTCTGGGCTCCAGGCTGAGAACGAGGCTGAATTATTACTGAGCTCATATACAGCACAG	356
Q	y		346	TTCACATGTCCTTGGGAACTGGGACCAAGGCACCGTCTCTGGTAGGCCCAAAGCCAACC	405
D	b		357	CATTATGATCTTTCGAACTGGAGACCAAGGTCACCGTCTTAGTATAGCCCAAAGCCAAACC	416
Q	y		406	CAGTGTCACTCTGTGTCCCGCCTCTCTTGAAGACTCCAAAGCCAAAGGCACACTAGT	465
D	b		417	CAGTGTCACTCTGTGTCCCGCCTCTCTTGAAGACTCCAAAGCCAAAGGCACACTAGT	476
Q	y		466	GTCGTCATCATGTAATCTTACCCGGAGCTGTGACAGTGGCCCTGGAAGGACGATGGAG	525
D	b		477	GTCGTCATCATGTAATCTTACCCGGAGCTGTGACAGTGGCCCTGGAAGGACGATGGAG	536
Q	y		526	CCCCGTGAAGCGGGAAGTGAGAACCAACCAACCTCCAAAACAGACCAACAAGTACGC	585
D	b		537	CCCCGTGAAGCGGGAAGTGAGAACCAACCAACCTCCAAAACAGACCAACAAGTACGC	596
Q	y		586	GGCCAGCAGCACTAAGTGAAGCTTGAAGCCCGAGCGACGTGAATGCCACAGAGCTACAGTG	645
D	b		597	GGCCAGCAGCACTAAGTGAAGCTTGAAGCCCGAGCGACGTGAATGCCACAGAGCTACAGTG	656
Q	y		646	CCAGGTACCGCATGAAGGAGACCCGCGATGTAAGACAGTGTGCC	691
D	b		657	CCAGGTACCGCATGAAGGAGACCCGCGATGTAAGACAGTGTGCC	701
RESULT 7					
HUMIGHPEPAL.			877 bp	mRNA	linear PRI 08-OCT-1993
LOCUS					
DEFINITION			Human (hybridoma H210) anti-hepatitis A immunoglobulin lambda chain variable region, constant region, complementarity-determining regions mRNA, complete cds.		
ACCESSION			M87790.1	GI:185363	
VERSION			M87790		
KEYWORDS					complementarity-determining region; constant region; immunoglobulin; lambda chain; light chain; variable region.
SOURCE					Homo sapiens (human)
ORGANISM					Homo sapiens
REFERENCE					Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS					1 (bases 1 to 877) Lewis,A.P., Parry,N., Peakman,T.C. and Crowe,J.S. Unpublished (1992)
JOURNAL					2 (bases 1 to 877) Lewis,A.P., Lemon,S.M., Barber,K.A., Murphy,P., Parry,N.R., Peakman,T.C., Sims,M.J., Wooten,J., and Crowe,J.S. Rescue, expression, and analysis of a neutralizing human anti-hepatitis A virus monoclonal antibody J. Immunol. 151 (5), 2829-2838 (1993)
FEATURES					93367243
COMMENT					8395549
FEATURES					Original source text: Homo sapiens cDNA to mRNA.
source					Location/Qualifiers
					1..877
					/organism="Homo sapiens"
					/mol_type="mRNA"
					/db_xref="taxon:9606"
					1.._31
					/note="putative"
					32..739
					/note="anti-hepatitis A; putative"
					/codon_start=1
					/product="immunoglobulin lambda-chain"
					/protein_id="AA02915.1"
					/db_xref="GI:185364"
					/translation="KAKALLLILFLIDODPSMAQSALTOPASVGSFGOSITSCGTG NNVGSTNLVSWIQHFGRAPFKMTIIVSRPGLSVNRFGSKXGNATSLISGLQAD
5' UTR					
CDS					

```

DEADYVCCSASVSATVFGGGTATLVGGPKAPSVTLPPSSEELQANKATLVCLIS
DPIFGATVAMKADSSPVKAGVETTTTPSKQSNKRYAASSTLSLTPEQMSHRSYSCQV
THEGSTVEKTVAPTECS"
sig_peptide
  /note="putative"
  32..88
mat_peptide
  /product="variable region; putative"
  422..736
mat_peptide
  /product="constant region; putative"
  155..196
gene
  /gene="CDR1"
  155..196
misc_feature
  /note="complementarity-determining region 1; putative"
  242..262
gene
  /gene="CDR2"
  242..262
misc_feature
  /note="complementarity-determining region 2; putative"
  359..388
gene
  /gene="CDR3"
  359..388
misc_feature
  /note="complementarity-determining region 3; putative"
  737..877
3'UTR
  842..847
polyA_signal
  /note="putative"
  853..858
ORIGIN
Query Match          40.5%; Score 597.4; DB 9; Length 877;
Best Local Similarity 92.5%; Pred. No. 1,2e-125;
Matches 639; Conservative 0; Mismatches 51; Indels 1; Gaps 1;

```

```

601 GAGCTGACGCCCGGACGAGCTGGAACTCCCAAGAACTACAGCTGCCAGGTCACGACATGA 660
634 GAGCCTGACGCGCTGAGCGAGTGGAACTCCCAAGAACTACAGCTGCCAGGTCACGACATGA 693
QY 661 AGGAGACACCGTGATGAAGAAGTGTGCC 691
DB 694 AGGAGACACCGTGA-GAAGACAGTGGCCCC 723
RESULT 8
AK130519
LOCUS
DEFINITION
  Homo sapiens CDNA FLJ27029 f1s, clone SLV05161, highly similar to
  IG lambda chain C regions.
ACCESSION
  AK130519
VERSION
  AK130519.1 GI:34527337
KEYWORDS
  oligo capping; f1s (full insert sequence).
SOURCE
  Homo sapiens
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1 Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S.,
    Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Horita, T., Watanabe, M.,
    Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M.,
    Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
    Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and
    Sugano, S.
    NEDO human cDNA sequencing project
    Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science,
    University of Tokyo, Laboratory of Genome Structure, Human Genome
    Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
    (E-mail: flicdn@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
    Fax: 81-3-5449-5416)
COMMENT
  NEDO human cDNA sequencing project supported by Ministry of
  Economy, Trade and Industry of Japan; cDNA full insert sequencing:
  Research Association for Biotechnology (RAB); cDNA library
  construction and 5'-end one pass sequencing: Institute of Medical
  Science, University of Tokyo, Laboratory of Genome Structure, Human
  Genome Center; 3'-end one pass sequencing: RAB; clone selection for
  full insert sequencing: RAB and Helix Research Institute.
FEATURES
  source
    1..1230
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="SLV05161"
    /tissue_type="salivary gland"
    /clone_id="SLV"
    /note="Cloning vector: pMB18SFL3"
  CDS
    42..755
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="BAC85368.1"
    /db_xref="GI:34527338"
    /translation="MAWALILTLITLGGTGSMAQSALTQPAVSQSGPSQISITISCTGS
    SSDVGVNFMFOOHGKAPKIVIVTVNRPSSVRSFGSKSGNTASLTISGLAE
    DEGVYCSSTYDKITTPSVVFGGKQVIVLQGPAPKAPVTLTPPSSEELQANKATLVCL
    ISDPIFGATVAMKADSSPVKAGVETTTTPSKQSNKRYAASSTLSLTPEQMSHRSYSCQ
    VTHEGSTVEKTVAPTECS"
ORIGIN
Query Match          40.5%; Score 591.4; DB 9; Length 1230;
Best Local Similarity 89.1%; Pred. No. 2,7e-124;
Matches 687; Conservative 0; Mismatches 71; Indels 13; Gaps 4;

```



Query Match 40.3%; Score 588.4; DB 9; Length 894;  
 Best Local Similarity 93.7%; Pred. No. 1,36-123;  
 Matches 650; Conservative 0; Mismatches 31; Indels 13; Gaps 3;

QY 1 GGCCTGGGCTGCTGATTCCTGACCCCTCCATCTCAGGAGCAGAGGCTCTGGGCCCCAGTC 60  
 Db 34 GGCCTGGGCTGCTGATTCCTGACCCCTCCATCTCAGGAGCAGAGGCTCTGGGCCCCAGTC 93

QY 61 TCCCTGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Db 94 TCCCTGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 153

QY 121 CACTGGAACGACGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 Db 154 CACTGGAACGACGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 213

QY 181 AGGCAAGCCCCCAACTCATCTATTATGAGTCAGTATCGGCTCCAGGCTTCTTA 240  
 Db 214 AGGCAAGCCCCCAACTCATCTATTATGAGTCAGTATCGGCTCCAGGCTTCTTA 273

QY 241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 Db 274 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 333

QY 301 TGAGGACGAGGCTGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357  
 Db 334 TGAGGACGAGGCTGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 393

QY 358 CGGAAGCTGGAGCCAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 417  
 Db 394 CGGCGAGGAGCCAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 453

QY 418 GTTCCCGCCCTCTGAGAGCTCCAGGCCAAGGCCAAGCTAGTGTGTGATGACG 477  
 Db 454 GTTCCCGCCCTCTGAGAGCTCCAGGCCAAGGCCAAGCTAGTGTGTGATGACG 504

QY 478 TGACTTCTACCCGGAGGCTGTGACAGTGGCTGAGAGGAGATGGAGCCCGCTCAAGGC 537  
 Db 505 TGACTTCTACCCGGAGGCTGTGACAGTGGCTGAGAGGAGATGGAGCCCGCTCAAGGC 564

QY 538 GGGAGTGGAGCCAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597  
 Db 565 GGGAGTGGAGCCAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624

QY 598 CCTGAGCTGAGCCCGAGAGGAGTGGAGTCCAGAGGATCAAGTGGAGTCAAGCA 657  
 Db 625 CCTGAGCTGAGCCCGAGAGGAGTGGAGTCCAGAGGATCAAGTGGAGTCAAGCA 684

QY 658 TGAAGGAGACCGTGTGATGAAGACAGTGTCC 691  
 Db 685 TGAAGGAGACCGTGTGATGAAGACAGTGTCC 717

RESULT 10  
 HSTGVL034 756 bp mRNA linear PRI 04-NOV-1994  
 LOCUS HSTGVL034  
 DEFINITION Human rearranged immunoglobulin lambda light chain mRNA.  
 ACCESSION X57824  
 VERSION X57824.1 GI:33747  
 KEYWORDS Ig lambda light chain; immunoglobulin.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 344 to 390)  
 AUTHORS Combarato, G. and Klobbeck, H.G.  
 TITLE V lambda and J lambda-C lambda gene segments of the human  
 immunoglobulin lambda light chain locus are separated by 14 kb and  
 rearrange by a deletion mechanism  
 JOURNAL Eur. J. Immunol. 21 (6), 1513-1522 (1991)  
 MEDLINE 91257162  
 PUBMED 1904362

REFERENCE 2 (bases 1 to 756)  
 AUTHORS Klobbeck, H.G.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-JAN-1991) H.G. Klobbeck, Inst fuer Physiologische  
 Chemie, Physiologische Biochemie und Zellbiologie der Universitaet  
 Muenchen, Shillerstr. 44, 8000 Muenchen 2, Germany  
 COMMENT for overlapping sequences see: X51754-55; J00252-54; M15641-42.

FEATURES  
 source  
 1..756  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /isolate="individual S"  
 /db\_xref="taxon:9606"  
 /chromosome="22"  
 /clone="C846"  
 /cell\_type="B-cell"  
 /tissue\_type="spleen"  
 /clone\_lib="phage library CS"  
 1..756  
 /gene="immunoglobulin lambda light chain"  
 1..756  
 /gene="immunoglobulin lambda light chain"  
 /note="cDNA"  
 /evidence="experimental"  
 1..708  
 /gene="immunoglobulin lambda light chain"  
 /codon\_start=1  
 /protein\_id="CAA40961.1"  
 /db\_xref="GI:33748"  
 /db\_xref="EMBL:CAA40961"  
 /translation="MAWTLFLTLTGTGSGMAQSALTQPAVSQSPQSLITISGCT  
 SSDIGAYFVWYQYPGKAPKLIPIVSGRPSGISRFSKSGKNTASLTISGLQAE  
 DEADYVCASVARSIVVGGSTVTVLQGPAPASVTLFPSPSEELCNKATVCLIS  
 DEPVAVTAMKADSSPVRAQVETTTTSKSNMNYAASVSLTPEQKSHRSYSCQV  
 THESVETKVAFTPCS"  
 1..57  
 /gene="immunoglobulin lambda light chain"  
 58..388  
 /gene="immunoglobulin lambda light chain"  
 /note="variable region  
 V(lambda)II"  
 352..388  
 /gene="immunoglobulin lambda light chain"  
 /note="J-segment"  
 389..708  
 /gene="immunoglobulin lambda light chain"  
 /note="constant region"

misc\_feature  
 352..388  
 /gene="immunoglobulin lambda light chain"  
 /note="J-segment"  
 389..708  
 /gene="immunoglobulin lambda light chain"  
 /note="constant region"

misc\_feature  
 352..388  
 /gene="immunoglobulin lambda light chain"  
 /note="J-segment"  
 389..708  
 /gene="immunoglobulin lambda light chain"  
 /note="constant region"

ORIGIN

Query Match 40.0%; Score 584.6; DB 9; Length 756;  
 Best Local Similarity 91.3%; Pred. No. 9,8e-123;  
 Matches 631; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

QY 1 GGCCTGGGCTGCTGATTCCTGACCCCTCCATCTCAGGAGCAGAGGCTCTGGGCCCCAGTC 60  
 Db 3 GGCCTGGGCTGCTGATTCCTGACCCCTCCATCTCAGGAGCAGAGGCTCTGGGCCCCAGTC 62

QY 61 TCCCTGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Db 63 TCCCTGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122

QY 121 CACTGGAACGACGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 Db 123 CACTGGAACGACGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 182

QY 181 AGGCAAGCCCCCAACTCATCTATTATGAGTCAGTATCGGCTCCAGGCTTCTTA 240  
 Db 183 AGGCAAGCCCCCAACTCATCTATTATGAGTCAGTATCGGCTCCAGGCTTCTTA 242

QY 241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 Db 243 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 302

QY	301	TGAGACGAGGCTGATATTACTGTGCTCATATCAAGAGTACTTTCATGCTTCG	360
Db	303	TGAGACGAGGCTGATATTATTGTGGCTTCATATCAAGAGTGCACGTAATCTTCG	362
QY	361	AACGTGAGCCAMAGTCAACCGTCTTAGTTCAGGCCAAGGCCAACCCCACTGTCACTGT	420
Db	363	CGAGAGGACCAAGGTACCGTCTTGTGTAGGCCAAGGCTGCCCTTCGTCACTGT	422
QY	421	CCGGCCTCTCTTGAGAGACTCCAMGCCACAGGCCCACTAGTGTGTATCACTGA	480
Db	423	CCGGCCTCTCTTGAGAGACTTCAAGCCACAGGCCCACTAGTGTGTCTCAATAGTA	482
QY	481	CTTCTACCCGGAGCGTGTACAGTGGCGCTGGAGCGAGATGGACCCCGCTCAAGCGGG	540
Db	483	CTTCTACCCGGAGCGGTGACAGTGGCGCTGGAGCGAGATGGACCCCGCTCAAGCGGG	542
QY	541	AGTGGAGACCCACCAAACTCTCCAAAACAGACACACACAGTACCGGCTCAGACACT	600
Db	543	AGTGGAGACCCACCACTCTCCAAAACAGACACACACAGTACCGGCTCAGACACT	602
QY	601	GAGCCTGACGCCCGAGCACTGGAAAGTCCCAAGAAAGTTACAGTTCAGAGTCAAGCATGA	660
Db	603	GAGCCTGACGCCCTGACCACTGGAAAGTCCCAAGAAAGTTACAGTTCAGAGTCAAGCATGA	662
QY	661	AGGAGACACCTGTGATGAAGACAGTGTCCC	691
Db	663	AGGAGACACCTGTGA-GAAGACAGTGTGCC	692

```

/protein_id="BAC01786.1"
/db_xref="IGI:21669533"
/translation="MKYLLPTPAAGLLILAACPAMASOALTQDASVSGSGASTISC
TGRSSDVGASVNRYSWQOCPGRTAKPLXIMIEVSRNPGSPDRFGSGSKGNFASLI156L
QADDDADYCCSSYTSSTSWVFGTGTATYVIGCKANPVTYLLFPSSSEELQANATLYC
LIDPDYGAATVAWAKDGGSPVAKGAVETDTSKQSNKYAASLITPPEWKSRRSRS
CQTHRESTVEKTVAPTECSAROSTPFVCEYQGSDDLPPQPNAGGSGSGS"
sig_peptide
1..66
/gene="IGI"
/note="pe1b signal peptide"

```

Query Match	40.0%	Score 584.4	IDB 9	Length 810
Best Local Similarity	95.0%	Pred. No. 1.1e-122		
Matches 614	Conservative 0	Mismatches 31	Indels 1	Gaps 1
Qy	46	GTCTGGGCCAGTGTGCCCCCTGACTAGTCTGCTCCGCTCGGTCTGGGCTCTCTCTGGACAGTC	105	
Db	57	GGCCATGAGCCAGAGTCTCCCTGACTAGCTGCTGCTCCGATCCGGGCTCTCTGGACAGTC	116	
Qy	106	GATACCATCTCCGCGACTGGAAACCGACATCACTGTTGGTGTATTAACATATGCTCTG	165	
Db	117	AGTACCATCTCTCGACTGGAAACCGACATGAGCTTGGTGTATTAACATATGCTCTCTG	176	
Qy	166	GTACCAACAGACCCCAAGCAAGCCCCCAACTCATCTATTATGAGGTACGTAATCGCC	225	
Db	177	GTACCAAGACCCCAAGCAAGCCCCCAACTCATCTATTATGAGGTACGTAATCGCC	236	
Qy	226	CTCAGGGGTTCTTAATGCTTCTCTGCTCCAGTCTGGCAACAGCCCTCCCTGACAT	285	
Db	237	CTCAGGGGTTCTTAATGCTTCTCTGCTCCAGTCTGGCAACAGCCCTCCCTGACAT	296	
Qy	286	CTCTGGGCTCAGAGCTGAGACGAGGCTGATTATCTGCTGCTCATATCAAGAGTAC	345	
Db	297	CTCTGGGCTCAGAGCTGAGACGAGGCTGATTATCTGAGCTCATATCAAGAGTAC	356	
Qy	346	TTCTCATGCTTTGGAACTGGGACCAAGTCAACGCTTATGTTAGTACGCCCAAGCC	405	
Db	357	CACTTGGGCTTTGGAACTGGGACCAAGTCAACGCTTATGTTAGTACGCCCAAGCC	416	
Qy	406	CACGTGACTCTGTTCGCCGCCCTCTCTGAGAGGCTCAAGCCCAAGGCAACACTAGT	465	
Db	417	CACGTGACTCTGTTCGCCGCCCTCTCTGAGAGGCTCAAGCCCAAGGCAACACTAGT	476	
Qy	466	GTGTCTGATCAGTACTTCTTACCCGGAGCTGTGACAGTGGCTTGAAGCAATGGCAG	525	
Db	477	GTGTCTGATCAGTACTTCTTACCCGGAGCTGTGACAGTGGCTTGAAGCAATGGCAG	536	
Qy	526	CCCCGTAAAGCCGGAGTGGAGACCAACCAACCTCCAAACAGAGCAACAAGTACGC	585	
Db	537	CCCCGTAAAGCCGGAGTGGAGACCAACCAACCTCCAAACAGAGCAACAAGTACGC	596	
Qy	586	GGCAGCAGCTTACTGAGCTGAGCGCCCGACAGTGAATCCACAGAACTACAGCTG	645	
Db	597	GGCAGCAGCTTACTGAGCTGAGCGCCCGACAGTGAATCCACAGAACTACAGCTG	656	
Qy	646	CCAGGTACGCGATGAAGGAGACACCGTGTATGAAGACAGTGTCTCC	691	
Db	657	CCAGGTACGCGATGAAGGAGACACCGTGTGAAGACAGTGTCTCC	701	
RESULT 12				
LOCUS	BC015833	915 bp	mRNA	1linear PRI 19-NOV-2003
DEFINITION	Homo sapiens cDNA clone MGC:27152 IMAGE:469160, complete cds.			
ACCESSION	BC015833			
VERSION	BC015833.1	GI:16198374		
KEYWORDS	MGC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 915)			



## AUTHORS

Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klasner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.N., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carrinck, P., Prange, C., Rana, S.S., Loguigliano, N.A., Peters, G.J.,  
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.N., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, U., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalls, D.E.,  
Schnerich, A., Schein, J.E., Jones, S.J., and Matra, M.A.,  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REMARK

## COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-rt@mail.nih.gov](mailto:cgapbs-rt@mail.nih.gov)  
Tissue Procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
Contact: [amadnan@systemsbio.org](mailto:amadnan@systemsbio.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRML Plate: 38 Row: K Column: 8  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis, Genomescan gene prediction.

## FEATURES

## source

## CDS

1..915  
location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:27152 IMAGE:4691630"  
/tissue\_type="lung"  
/clone\_id="NIH\_MGC\_77"  
/lab\_host="DH10B"  
/note="vector: pDNR-LIB"  
43..750  
/codon\_start=1  
/product="Unknown (protein for MGC:27152)"  
/protein\_id="AAH1833.1"  
/db\_xref="GI:16198375"  
/translation="MAMALLLTLLTDTGWSAQSALTQPAVSQSGSGGTTTCTGT  
STDVSHLSVSNYQHPGKAPKFLIFSGSKPSVSNRPSGSKGNASLTISGLQRE  
DEADYVCSYVSGTGVGGTKLTLVQPPAASVTLFPSSSELQANXATVCLIS  
DFYPAVTVMKADSSPVKAGVETTPESKSNKXAASSYSLTPEQMKSHRSYSCV  
THEGSTVKTVAPECS"  
127..378  
/note="IGV: Region: Immunoglobulin domain variable region  
(V) subfamily"  
/db\_xref="CDD:cd00099"  
445..729  
misc\_feature  
misc\_feature  
misc\_feature

## ORIGIN

Query Match 39.8%; Score 581.4; DB 9; Length 915;  
Best local similarity 91.0%; Pred. No. 5,381-122;  
Matches 629; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

/note="IGV: Region: Immunoglobulin domain constant region  
subfamily"  
/db\_xref="CDD:cd00098"

1 GGCTGGGCTCTGCTCTATTTCTACCTCTCTACCTAGGACAGAGGTCCTGGGCCAGATC 60  
45 GGCTGGGCTCTGCTCTGCT 104  
61 TGCCCTGACTCAGTCTGCT 120  
105 TGCCCTGACTCAGTCT 164  
121 CACTGGAACCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
165 CACTGGAACCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 224  
181 AGGCAAGAGCCCAAGCTATCATTTATGAGGTCAGTAACGAGCCCTCAGGGGTTCTAA 240  
225 AGGCAAGAGCCCAAGCTATTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 284  
241 TGCCCTGCTGCT 300  
285 TGCCCTGCTGCT 344  
301 TGAGGACGAGGCTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 360  
345 TGAGGACGAGGCTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 404  
361 AACTGGAGCAAGAGTACCGCT 420  
405 CGAGGAGGAGGAGTACCGCT 464  
421 CCGGCT 480  
465 CCGGCT 524  
481 CTTCT 540  
525 CTTCT 584  
541 AGTGAAGACCAACCAACCT 600  
585 AGTGAAGACCAACCAACCT 644  
601 GAGCTGAGCGCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 660  
645 GAGCTGAGCGCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 704  
661 AGGAGACCAACCT 691  
705 AGGAGACCAACCT 734

## RESULT 13

## HSIGVL001

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

Human rearranged immunoglobulin lambda light chain mRNA.  
X57801.1 GI:33699  
Ig lambda light chain; immunoglobulin.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 372 to 418)  
c-myc, c-myc, and c-myc.  
V lambda and V lambda-C lambda gene segments of the human  
immunoglobulin lambda light chain locus are separated by 14 kb and

REARRANGE BY A DELETION MECHANISM  
 FOR J. Immunol. 21 (6), 1513-1522 (1991)  
 MEDLINE  
 PUBMED  
 1904362  
 2 (bases 1 to 783)  
 Klobeck, H.G.  
 Direct Submission

COMMENT  
 Submitted (31-JUN-1991) H.G. Klobeck, Inst fuer Physiologische  
 Chemie, Physiologische Biochemie und Zellbiologie der Universitaet  
 Muenchen, Shillerstr. 44, 8000 Muenchen 2, Germany  
 for overlapping sequences see: X51754-55; J00252-54; M15641-42.  
 FEATURES  
 source  
 1..783  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="C2132"  
 /cell\_line="GM2132"  
 /cell\_type="B-cell"  
 /clone\_1ib="phage library c2132"  
 1..783  
 /gene="immunoglobulin lambda light chain"  
 1..783  
 /gene="immunoglobulin lambda light chain"  
 /note="cDNA"  
 /evidence="experimental"

## CDS

23..736  
 /gene="immunoglobulin lambda light chain"  
 /codon\_start=1  
 /protein\_id="CAA0939.1"  
 /db\_xref="GI:33700"  
 /translation="VAMVALLLTLTGCGSNASQATNPASVNGSPGQILITSGTP  
 SPSIDGYQYISWYQCHPKAPKLLIYDVXKRSQSVSNRPSGSKSNTSLITSGIQAE  
 DEADYICSSIRGSATFEVFGGKTVLQPKAPSVYLPFPSEELQANKATIVCL  
 ISDFYGAIVAMKADSPVKAQVETTPSKQSNKRYAASVYLSLPQWMSHRSYSC  
 QVTHSGSTVEKTVAPTECS"  
 23..79  
 /gene="immunoglobulin lambda light chain"  
 80..416  
 /gene="immunoglobulin lambda light chain"  
 /note="variable region  
 V(lambda)II"  
 380..416  
 /gene="immunoglobulin lambda light chain"  
 /note="J-segment"  
 417..736  
 /gene="immunoglobulin lambda light chain"  
 /note="constant region"

## sig\_peptide

80..416  
 /gene="immunoglobulin lambda light chain"  
 /note="variable region  
 V(lambda)II"  
 380..416  
 /gene="immunoglobulin lambda light chain"  
 /note="J-segment"  
 417..736  
 /gene="immunoglobulin lambda light chain"  
 /note="constant region"

## misc\_feature

417..736  
 /gene="immunoglobulin lambda light chain"  
 /note="constant region"

## ORIGIN

Query Match 39.1%; Score 570.2; DB 9; Length 783;  
 Best Local Similarity 90.7%; Pred. No. 1.9e-119;  
 Matches 632; Conservative 0; Mismatches 58; Indels 7; Gaps 2;  
 QY 1 GGCCTGGGCTCTCTATTTCTCAACCTCTCACTCAAGGCAAGGGTCTGGGCCCAAGTC 60  
 DB 25 GGCCTGGGCTCTCTCTCTCTCAACCTCTCACTCAAGGCAAGGGTCTGGGCCCAAGTC 84  
 QY 61 TGCCCTGACTCACTGACCTCCGCTGCTGGGCTCTCTGACAGTGGATCACATCTCTG 120  
 DB 85 TGCCCTGACTCACTGACCTCCGCTGCTGGGCTCTCTGACAGTGGATCACATCTCTG 144  
 QY 121 CACTGGAACCAAGAGTCAAGTGGTGTATTAATCTCTCTCTGATCAACAGACACCC 180  
 DB 145 CACTGGAACCAAGAGTCAAGTGGTGTATTAATCTCTCTGATCAACAGACACCC 204  
 QY 181 AGGCAAG 240  
 DB 205 AGGCAAG 264  
 QY 241 TCGCTTCTCTGCTCAAGTCTGGCAACAGGCTCTCCCTGACACATCTCTGGGCTCCAGGC 300

DB 265 TCGCTTCTCTGCTCAAGTCTGGCAACAGGCTCTCCCTGACACATCTCTGGGCTCCAGGC 324  
 QY 301 TGAAG 354  
 DB 325 TGAAG 384  
 QY 355 CTTTGGAGTGGGACCAAGGTCAACCGTCTTGAAGTCAAGCCCAAGCCCACTGTCAAC 414  
 DB 385 GTTCGGGAGGAGGACCAAGGTGACCGTCTTGAAGTCAAGCCCAAGCCCACTGTCAAC 444  
 QY 415 TCTGTCCTCCGCTCTCTGAGAGAGTCCAGGCAACAGGCGACACTGATGTCTGAT 474  
 DB 445 TCTGTCCTCCGCTCTCTGAGAGAGTCCAGGCAACAGGCGACACTGATGTCTGAT 504  
 QY 475 CAGTGAATTTACCCGAGAGCTGTGACAGTGGCTGGAGGAGAGAGAGAGAGAGAGAGAG 534  
 DB 505 AAGTGAATTTACCCGAGAGCTGTGACAGTGGCTGGAGGAGAGAGAGAGAGAGAGAGAGAG 564  
 QY 535 GCGGAGTGGAGACCAACCGTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594  
 DB 565 GCGGAGTGGAGACCAACCGTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 624  
 QY 595 CTACCTGAGCTGACCGCCGAGAGTGAAGTCCACAGAGCTTACAGCTGCGAGGTCAAC 654  
 DB 625 CTACCTGAGCTGACCGCTGAGAGAGTGAAGTCCACAGAGCTTACAGCTGCGAGGTCAAC 684  
 QY 655 GCATGAAGGAGACCGCTGATGAAGACAGTGTGCTCC 691  
 DB 685 GCATGAAGGAGACCGCTGATGAAGACAGTGTGCTCC 720

RESULT 14  
 AB064155  
 LOCUS  
 AB064155 806 bp mRNA linear PRI 02-JUL-2002  
 DEFINITION  
 Homo sapiens IGL mRNA for immunoglobulin lambda light chain V(LJ  
 region, partial cds, clone:115.  
 AB064155  
 VERSION  
 AB064155.1 GI:21669516  
 KEYWORDS  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1  
 Akahori, Y., Iba, Y., Morino, K., Shinohara, M., Hirono, Y., Kakita, M.,  
 Suzuki, K., Torii, H., Uka, Y., Honda, T., Katsumi, H., Okada, J.,  
 Miura, K., and Kurosawa, Y.  
 Construction and characterization of antibody libraries: isolation  
 of therapeutic human antibodies and application to functional  
 genomics

## JOURNAL

Unpublished  
 2 (bases 1 to 806)  
 Kurosawa, Y.  
 Direct Submission  
 Submitted (25-JUN-2001) Yoshikazu Kurosawa, Institute for  
 Comprehensive Medical Science, Fujita Health University,  
 Kutsukake-cho, Toyosake 470-1192, Japan  
 (E-mail: kurosawa@fujita-hu.ac.jp, Tel:81-562-93-9387)  
 Please visit our web site  
 URL: http://www.fujita-hu.ac.jp/immunity/.

## COMMENT

## FEATURES

source  
 1..806  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="115"  
 /clone\_1ib="AIMS4"  
 /note="mixture of tissues: tonsils, umbilical cords,  
 peripheral blood and bone marrow"  
 1..806  
 /gene="IGL"  
 1..806  
 CDS

Query Match	39.0%	Score 570;	DB 9;	Length 806;
Best Local Similarity	93.7%	Pred. No. 2.1e-119;		
Matches 605;	Conservative	0;	Mismatches 40;	Indels 1;
				Gaps 1;

RESULT 15  
AB064163

LOCUS	AB064163	810 bp	mRNA	linear	PRI 02-JUL-2002
DEFINITION	Homo sapiens IGL mRNA for immunoglobulin lambda light chain VJ1 region, partial cds, clone:123.				
ACCESSION	AB064163				
VERSION	AB064163.1	GI:21669532			
KEYWORDS	Homo sapiens (human)				
SOURCE					

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cdone="L23"
/clone_id="AIMS4"
/note="mixture of tissues:tonsils, umbilical cords,
peripheral blood and bone marrow"
1..810
gene

```

```

CDS
/gene="IGL"
1..>810
/gene="IGL"
/codon_start=1
/product="immunoglobulin lambda light chain VLJ region"
/protein_id="BAC01791.1"
/db_xref="NCI:2166953"
/translation="MKYILPTAAAGLILAAOPAMASALTOPASVSGSGPGSTTSC
TGRSSVVGGINVSYNQCHGKAPKMTIDVSRPFGSGKSRPFGSKGNATSLISLI
QAEDADYICSTSSSSWFGGKGLTILGQPKAASVILFPSSSELDANKATLVLC
LIDFPYGAIVAMKADSSVFKAGVETLTPPSKSNKKPAASVLSLILPEQMKHSYS
CQYTHGSTEYKTVAPTECSARQSTPFVCEYIQGSSDLPPQPVNAGGSGGG"
sig_peptide
1..66
/gene="IGL"
/note="pe1b signal peptide"

ORIGIN
Query Match 39.0%; Score 570; DB 9; Length 810;
Best Local Similarity 93.7%; Pred. No. 2.1e-119;
Matches 605; Conservative 0; Mismatches 40; Indels 1; Gaps 1;

```

QY	46	GTCTGGGCCAGTCTGCCCTGACTAGTGTGACTCTGGGCTCTCTGGACAGTC	105
Db	57	GGCCATGGCCCACTGTGCTCCCTGACTCAGCTGCCTCCGTCTCGGGCTCCTGGACAGTC	116
QY	106	GATCACCATTTCTGTGACTGGAAACAGCAGTCAAGTTGGTGTATATCATATGCTCTGT	165
Db	117	GATCACCATTTCTGTGACTGGAAACAGCAGTCAAGTTGGTGTATATCATATGCTCTGT	176
QY	166	GTACCAACAGCACCAGCAAGACCCCAACTCATCTATTAGAGGTCAATATCGGCC	225
Db	177	GTACCAACAGCACCAGCAAGACCCCAACTCATCTATTAGAGGTCAATATCGGCC	236
QY	226	CTCAGGGGTTTCTTAATGCTTCTCTGTGCTTCAAGTCTGGCAACAGGCTCTCTGACCAT	285
Db	237	CTCAGGGGTTTCTTAATGCTTCTCTGTGCTTCAAGTCTGGCAACAGGCTCTCTGACCAT	296
QY	286	CTCTGGGCTCAGGCTGAGACGAGGCTGATTATTAATCTGTGCTCATATCAAGAAATAC	345
Db	297	CTCTGGGCTCAGGCTGAGACGAGGCTGATTATTAATCTGAGCTCATATCAAGAAATAC	356
QY	346	TTCTCATGTCTTCGAACTGGGAACCAAGTCAACCGTCTTAGTCAAGCCCAAGCCAAACC	405

```
Db 357 CTCTTGAGTGTTCGGCGAGGAGCAAGCTGACCGTCTAGGTCAGCCCAAGGCTGCCCC 416
QY 406 CACTGTCACTCTGTTCCTCCGCCCTCTCTGAGAGCTCCAAGCCAAAGGCCACACTAGT 465
Db 417 CTGGTCACTGTGTCCGCCCTCTCTGAGAGCTTCAAGCCAAAGGCCACACTGTGT 476
QY 466 GTGTCTGATCAGTGACTTCTTACCCGAGAGCTGTGACAGTGCCTGAAAGGCAGATGGCAG 525
Db 477 GTGTCTCATTAAGTACTTCTTACCCGAGAGCCGTGACAGTGGCTTGAAGGCAGATAGCAG 536
QY 526 CCCCGTCAAGGCGGAGTGAAGCAACCAACCTTCCAAACAGAGCAACACAGTACGC 585
Db 537 CCCCGTCAAGGCGGAGTGAAGCAACCAACCTTCCAAACAGAGCAACACAGTACGC 596
QY 586 GGCACAGAGCTACTGAGCCTGACGCCGAGCAGTGAAGTCCACAGAAAGCTACAGCTG 645
Db 597 GGCACAGAGCTACTGAGCCTGAGCCTGAGCAGTGAAGTCCACAGAAAGCTACAGCTG 656
QY 646 CCAAGTCAAGCAGTGAAGGAGCAACCGTGAATGAAGCAAGTGTGCC 691
Db 657 CCAAGTCAAGCAGTGAAGGAGCAACCGTGAAGTGAAGCAAGTGTGCC 701
```

Search completed: April 26, 2004, 18:46:55  
Job time : 5606.61 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 14:04:56 ; Search time 555.73 Seconds  
(without alignments)  
1160.755 Million cell updates/sec

Title: US-10-076-747-8

Perfect score: 1660  
Sequence: 1 ggcctggcctctgtatcc.....ggcagcagagccacgcagct 1460

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373663 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*\n2: Geneseq1990s:\*\n3: Geneseq2000s:\*\n4: Geneseq2001as:\*\n5: Geneseq2001bs:\*\n6: Geneseq2002s:\*\n7: Geneseq2003as:\*\n8: Geneseq2003bs:\*\n9: Geneseq2003cs:\*\n10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1459	99.9	1460	7	ABX92194 Human ova
2	809.6	55.5	957	7	ABX92193 Human ova
3	616.8	42.2	1636	5	ABV22585 Human pro
4	616.8	42.2	1636	5	ABV22585 Human pro
5	607	41.6	891	4	AAC65528 Human imm
6	597.4	40.9	902	2	AAC35100 Antibody
7	597.4	40.9	1480	8	ADAI1028 Human CDN
8	589.8	40.8	763	5	AAS83485 DNA encod
9	589.8	40.4	763	5	AAS83485 DNA encod
10	584.6	40.0	756	5	AAS83477 DNA encod
11	584.6	40.0	1636	5	ABV22585 Human pro
12	584.6	40.0	1636	5	ABV22585 Human pro
13	581.2	39.8	2667	9	AAD59189 Human rea
14	578.2	39.6	889	5	AAS77073 DNA encod
15	578.2	39.4	883	5	AAD59186 Human (hy
16	570.2	39.1	783	5	AAS83483 DNA encod
17	564.8	38.7	866	5	AAS87037 DNA encod
18	562.2	38.5	876	5	AAS83478 DNA encod
19	555.6	38.1	811	9	AAD59189 Human rea
20	529.4	36.3	726	7	ABX12863 DNA encod
21	511.8	35.1	2112	6	ABQ75369 Human lun
22	508	34.8	777	6	AAS77072 DNA encod
23	498.4	34.1	935	4	AAC65525 Human imm

24	497.2	34.1	605	6	ABK29738 Colon ade
25	488.6	33.5	9532	7	ABZ37360 MALTA3 nu
26	482	33.0	870	6	AAU50812 Human can
27	480.6	32.9	543	6	ABK38932 CDNA enco
28	480.6	32.9	543	7	ACA11261 Human lun
29	480.6	32.9	543	7	ACA02447 Lung can
30	480.4	32.9	884	2	AAQ03609 Sequence
31	477.4	32.7	1845	6	AB149526 Plasmid s
32	473.4	32.4	930	4	AAS22541 Human CDN
33	473	32.4	848	6	AAD59183 Human IG
34	468.2	32.1	915	6	AB165478 Lung can
35	468.2	32.1	915	6	ABK64815 Human ben
36	468.2	32.1	915	6	ABN97248 Gene #374
37	468.2	32.1	915	7	ACA64802 Human IG
38	468.2	32.1	964	6	ABQ93374 Human CDN
39	466.6	32.0	762	5	AAC84209 Plasmid G
40	465.2	31.9	5679	5	AAC84207 Plasmid G
41	463.4	31.7	951	5	AAS90539 DNA encod
42	462.8	31.7	888	5	AAS9184 Human IG
43	459.6	31.5	810	5	AAS87270 DNA encod
44	458	31.4	768	2	AAK06954 Monoclon
45	458	31.4	768	2	AAK06953 Monoclon

## ALIGNMENTS

RESULT 1	ABX92194	ABX92194 standard, cDNA, 1460 BP.
AC	XX	ABX92194;
DT	08-MAY-2003	(first entry)
XX	XX	Human ovarian specific nucleic acid DEX0310_8.
XX	XX	Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;
KW	non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;	
KM	gynecological.	
XX	XX	Homo sapiens.
OS	XX	WO200292785-A2.
PN	XX	21-NOV-2002.
PD	XX	13-FEB-2002; 2002WC-US022271.
PF	XX	13-FEB-2001; 2001US-0268290P.
PR	XX	15-FEB-2001; 2001US-0268834P.
PA	XX	(DIAD-) DIADEXUS INC.
PI	XX	Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferty R;
PI	XX	Sun Y, Liu C;
XX	XX	WPI, 2003-120677/11.
DR	XX	P-FSDB; ABU61023.
XX	XX	New isolated OSNA nucleic acid and encoded polypeptide, useful for
PT	XX	identifying, diagnosing, monitoring, staging, imaging and treating
PT	XX	ovarian cancer and non-cancerous diseases in ovarian tissues.
PS	XX	Claim 1, Page 146, 224pp; English.
XX	XX	The invention relates to a new isolated nucleic acid termed ovarian
CC	XX	specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that
CC	XX	encodes any of 53 fully defined protein sequences appearing as ABU61018-
CC	XX	ABU61070 (termed ovarian specific proteins, OSP); (b) any of 76 fully
CC	XX	defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a
CC	XX	sequence having at least 60% sequence identity to the nucleic acid
CC	XX	molecule of (a) or (b). Also included are a method for determining the



DR MPI: 2003-120677/11.  
 DR P-P8DB; ABU61022.  
 XX New isolated OSNA nucleic acid and encoded polypeptide, useful for  
 PT identifying, diagnosing, monitoring, staging, imaging and treating  
 PT ovarian cancer and non-cancerous diseases in ovarian tissues.  
 XX  
 PS Claim 1; Page 145; 224pp; English.  
 XX  
 CC The invention relates to a new isolated nucleic acid, termed ovarian  
 CC specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that  
 CC encodes any of 53 fully defined protein sequences appearing as ABU61018-  
 CC ABU61070 (termed ovarian specific proteins, OSP); (b) any of 76 fully  
 CC defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a  
 CC sequence having at least 60% sequence identity to the nucleic acid  
 CC molecule of (a) or (b). Also included are a method for determining the  
 CC presence of an ovary specific nucleic acid (OSNA) in a sample, a vector  
 CC comprising an OSNA, a host cell comprising the vector, an isolated OSP  
 CC polypeptide, an anti-OSP antibody or fragment, a method for determining  
 CC the presence of an ovary specific protein in a sample and a vaccine  
 CC comprising an OSP or OSNA. The methods and compositions of the present  
 CC invention are useful for identifying, diagnosing, monitoring, staging,  
 CC imaging and treating ovarian cancer and non-cancerous disease in ovary  
 CC tissue. The present sequence is an OSNA of the invention  
 CC  
 SQ Sequence 957 BP; 238 A; 290 C; 288 G; 140 T; 0 U; 1 Other;  
 Query Match 55.5%; Score 809.6; DB 7; Length 957;  
 Best Local Similarity 96.5%; Pred. No. 1.3e-180;  
 Matches 864; Conservative 0; Mismatches 19; Indels 12; Gaps 3;  
 QY 578 AAGTACGGCGCAGCAGCTACCTGAGCTGAGCCGCGGAGTGAAGTCCACAG 633  
 DB 63 AGGTACCGCGTCAGAGCTATCTGAGCTGAGCGCTGATGAGTGAAGTCCACAG 122  
 QY 634 AAGTACAGCTGCCAGCTAC-----GCATGAAGGAG-CACCGTGAATGAAGACGT 685  
 DB 123 AAGCTACAGCCTGACAGAGTCTTAGCTGATGAAGGAGACACCGGTGAAGACGT 182  
 QY 686 GATCCCTACAGAAATGTTCTATAGTCTCTTAACAGCCTACGCCCACTACCAATGAG 745  
 DB 183 GGTCCCTACAGAAATGTTCTATAGTCTCTTAACAGCCTACGCCCACTACCAATGAG 242  
 QY 746 GAGACTACAGCCTGACAGGATCCCAAGGAGAGGGGCTCTCTCATCAACACACAG 805  
 DB 243 GAGACTACAGCCTGACAGGATCCCAAGGAGAGGGGCTCTCTCATCAACACACAG 302  
 QY 806 GCGGAGTAAAGCCCTTATCATCGGCAATGTCAGCGTCAATAAGCGCTCAATAA 865  
 DB 303 GCGGAGTAAAGCCCTTATCATCGGCAATGTCAGCGTCAATAAGCGCTCAATAA 362  
 QY 866 GATATTCGACCTGGTGAATGTCCTACATACATACAAAGACATACCTCAACTACA 925  
 DB 363 GATATTCGACCTGGTGAATGTCCTACATACATACAAAGACATACCTCAACTACA 422  
 QY 926 CGGAGTCATGATACCGGTCGCGGCGGGAATGCACTCCCTCTCGGCGCTCAAT 985  
 DB 423 CGGAGTCATGATACCGGTCGCGGCGGGAATGCACTCCCTCTCGGCGCTCAAT 482  
 QY 986 GATGACTATGATACACCGGTCGCGGCGGCAANTATGATGATATTCGCCCGGG 1045  
 DB 483 GATGACTATGATACACCGGTCGCGGCGGCAANTATGATGATATTCGCCCGGG 542  
 QY 1046 ATCAAGACCATGATTTCCCGCGGTGATGATATCGGTAAGCGGCAATCAATT 1105  
 DB 543 ATCAAGACCATGATTTCCCGCGGTGATGATATCGGTAAGCGGCAATCAATT 602  
 QY 1106 CGCCCGGAGGCTACAGACCTAAAGTGAACCGGCGAGCGGCGAAGACAGGCTTTCG 1165  
 DB 603 CGCCCGGAGGCTACAGACCTAAAGTGAACCGGCGAGCGGCGAAGACAGGCTTTCG 662  
 QY 1166 ACGGATGCGAAACACAGAGGCGCGACAGCCGCGAGCTCCGCGGAGAGGCCCA 1225

DB 663 ACGGATGCGAAACACAGAGGCGCGACAGCCGCGAGGAGACCCCGCGGAGAGCCCA 722  
 QY 1226 CTGCGGACAGGCGGAGCGGAGGAGACATACAGAGACGCGAGCCACACCGCGCTGCA 1285  
 DB 723 CTGCGGACAGGCGGAGCGGAGGAGACATACAGAGACGCGAGCCACACCGCGCTGCA 782  
 QY 1286 TGAGACGAGACGCGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1345  
 DB 783 TGAGACGAGACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 842  
 QY 1346 AGCCACACGCGCTGAGGCGCTGCGGACAGCGCGCGGAGGAGGAGGAGGAGGAGGAG 1405  
 DB 843 AGCCACACGCGCTGAGGCGCTGCGGACAGCGCGCGGAGGAGGAGGAGGAGGAGGAG 902  
 QY 1406 CGGAGGAGACCGGACGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1460  
 DB 903 CGGAGGAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 957  
 RESULT 3  
 ID ABV22585  
 XX ABV22585 standard; cDNA; 1636 BP.  
 AC ABV22585;  
 XX  
 DT 13-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 22576.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US005171.  
 XX  
 PR 17-FEB-2000; 2000US-0189319P.  
 XX  
 PR 16-MAR-2000; 2000US-0189862P.  
 XX  
 PR 25-MAY-2000; 2000US-0207454P.  
 XX  
 PR 09-JUN-2000; 2000US-0211314P.  
 XX  
 PR 18-JUL-2000; 2000US-0219007P.  
 XX  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 XX Schlegel R, Endege WO, Monahan JE;  
 XX  
 DR MPI; 2001-662795/76.  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1; Page 3949; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 1636 BP; 345 A; 470 C; 479 G; 342 T; 0 U; 0 Other;

Query Match 42.2%; Score 616.8; DB 5; Length 1636;  
 Best Local Similarity 95.7%; Pred. No. 3,3e-135;  
 Matches 668; Conservative 0; Mismatches 22; Indels 8; Gaps 3;

1 GGCCCTGGGCTGCTGATTCCTCAACCCCTCACTCAGGGGACAGAGGCTCTGGGCCAGTC 60  
 62 GGCCCTGGGCTGCTGATTCCTCAACCCCTCACTCAGGGGACAGAGGCTCTGGGCCAGTC 121  
 61 TGCCCTGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 122 TGCCCTGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181  
 121 CACTGGAACGACGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 179  
 182 CACTGGAACGACGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241  
 180 CAGGCAAGCCCCCAACTCATATTATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 239  
 242 CAGGCAAGCCCCCAACTCATATTATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 301  
 240 ATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 239  
 302 ATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361  
 300 CTGAGGACGAGGCTGATTTATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 353  
 362 CTGAGGACGAGGCTGATTTATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421  
 354 TCTTGGAACTGGGACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 413  
 422 TCTTGGAACTGGGACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 481  
 414 CTCTGTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 473  
 482 CTCTGTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541  
 474 TCACTGAACTTACCCGCGGAGCTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 533  
 542 TCACTGAACTTACCCGCGGAGCTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 601  
 534 AGGGGGGAGTGAACCAACCAACCCCTCCAAACGAGCAACCAACCAACCAACCAACCAACCA 593  
 602 AGGGGGGAGTGAACCAACCAACCCCTCCAAACGAGCAACCAACCAACCAACCAACCAACCA 661  
 594 GCTACCTGAGCTGACGCGCGGAGCTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 653  
 662 GCTACCTGAGCTGACGCGCGGAGCTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 721  
 654 CGCATGAAGGAGCAACCGTGAATGAAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 691  
 722 CGCATGAAGGAGCAACCGTGAATGAAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 758

RESULT 4  
 ABV28405  
 ID ABV28405 standard; cDNA; 1636 BP.  
 XX  
 AC ABV28405;  
 XX  
 DT 16-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 28395.  
 XX  
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 XX pharmanomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX

20-FEB-2001; 2001MO-US005171.  
 XX  
 XX 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PI Schlegel R, Endege WO, Monahan JE;  
 DR MPI; 2001-662795/76.  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1; Page 5921; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC / (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 CC  
 SQ Sequence 1636 BP; 345 A; 470 C; 479 G; 342 T; 0 U; 0 Other;  
 Query Match 42.2%; Score 616.8; DB 5; Length 1636;  
 Best Local Similarity 95.7%; Pred. No. 3,3e-135;  
 Matches 668; Conservative 0; Mismatches 22; Indels 8; Gaps 3;

1 GGCCCTGGGCTGCTGATTCCTCAACCCCTCACTCAGGGGACAGAGGCTCTGGGCCAGTC 60  
 62 GGCCCTGGGCTGCTGATTCCTCAACCCCTCACTCAGGGGACAGAGGCTCTGGGCCAGTC 121  
 61 TGCCCTGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 122 TGCCCTGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181  
 121 CACTGGAACGACGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 179  
 182 CACTGGAACGACGATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241  
 180 CAGGCAAGCCCCCAACTCATATTATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 239  
 242 CAGGCAAGCCCCCAACTCATATTATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 301  
 240 ATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 239  
 302 ATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361  
 300 CTGAGGACGAGGCTGATTTATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 353  
 362 CTGAGGACGAGGCTGATTTATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421  
 354 TCTTGGAACTGGGACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 413  
 422 TCTTGGAACTGGGACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 481  
 414 CTCTGTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 473  
 482 CTCTGTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541  
 474 TCACTGAACTTACCCGCGGAGCTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 533





QY	181	AGGCAAGAGCCCCCAATCATCATTTATGAGTGAAGTATGGCCCTCAGGGGTTTCTAA	240
Db	214	AGGCAAGAGCCCCCAATCATCATTTATGAGTGAAGTATGGCCCTCAGGGGTTTCTAA	273
QY	241	TGCGCTTCTGCGTCCAAAGTCTGGCAACAGCGCTCCCTGAACATCTTCTGGGCTCCAGGC	300
Db	274	TTCCTTCTCTGGCTCCAAAGTCTGGCAACAGCGCTCCCTGAACATCTTCTGGGCTCCAGGC	333
QY	301	TGAGACAGAGGCTGATTATTACTGTGCTCATATACAAGAAAGTACTTCTCATGCTTCGG	360
Db	334	TGAGACAGAGGCTGATTATTACTGTGCTCATATACAAGAAAGTACTTCTCATGCTTCGG	393
QY	361	AACTGGAGCCAAAGTCAACCGTCTTAGGTCAAGCCCAAGGCCAACCCCACTGTCACTCTGTT	420
Db	394	CGAGAGGAGCCAAATGACCCGTCTAGGTCAAGCCCAAGGCCAACCCCTTCGCTCACTCTGTT	453
QY	421	CCGGCCCTCCTCTGAGAGAGCTCCAAAGCCCAAGGCCCACTAGTGTGCTGATCAAGTGA	480
Db	454	CCGGCCCTCCTCTGAGAGAGCTCCAAAGCCCAAGGCCCACTAGTGTGCTGATCAAGTGA	513
QY	481	CTTCTACCCGGAGAGCTGTGACAGTGGCTTGAAGGACAGATGGACGCCCTTCAAAGGCGGG	540
Db	514	CTTCTACCCGGAGAGCTGTGACAGTGGCTTGAAGGACAGATGGACGCCCTTCAAAGGCGGG	573
QY	541	AGTGGAGACCAACCAACCTCTCCAAACAGACAAACAAGTACCGCGGCCAGCGCTACTCT	600
Db	574	AGTGGAGACCAACCAACCTCTCCAAACAGACAAACAAGTACCGCGGCCAGCGCTACTCT	633
QY	601	GACCTCGAGCGCCGAGCAGTGGAAAGTCCCAAGAAAGTACAGCTCCAGGTCACGCATGA	668
Db	634	GACCTCGAGCGCCGAGCAGTGGAAAGTCCCAAGAAAGTACAGCTCCAGGTCACGCATGA	693
QY	661	AGGAGACACCCGTGATGAAAGACAGTGTCCC	691
Db	694	AGGAGACACCCGTGAG-GAAGACAGTGGCCCC	723
RESULT 7			
ADAl1028/c			
ID	ADAl1028 standard; cDNA; 1480 BP.		
XX	ADAl1028;		
XX	AC		
XX	DT 06-NOV-2003 (first entry)		
XX	DE Human cDNA differentially expressed in colon cancer #101.		
XX	KW ss; differential expression; colon cancer; cancer; human.		
XX	OS Homo sapiens.		
XX	PN US2002160382-A1.		
XX	PD 31-OCT-2002.		
XX	PF 11-OCT-2001; 2001US-00981351.		
XX	PR 11-OCT-2000; 2000US-0239841P.		
XX	PA (LASEK/) LASEK A W.		
XX	PA (JONE/) JONES D A.		
XX	PI Lasek AW, Jones DA;		
XX	PI WPI; 2003-265756/26.		
XX	DR		
XX	PT New combination comprising cDNAs that are differentially expressed in		
XX	PT colon disorder, useful for diagnosing, treating, staging or monitoring		
XX	PT treatment for colon cancers.		
XX	PS Claim 1; SEQ ID NO 146; 231p; English.		
XX	XX		
CC	The invention relates to a combination comprising cDNAs that are		









```

QY 129 CCAAGACGACGTTGGTGGTATATATATATCTCTCTGTTACCAACAGACCCAGGCAAG 188
DB 1432 CCAAGACGACGTTGGTGGTATATATATATCTCTCTGTTACCAACAGACCCAGGCAAG 1373
QY 189 CCCCCAACTCATATTTATGAGTCAAGTATCGGCTCAGGGGTTCTTAATGCTTCT 248
DB 1372 CCCCCAACTCATATTTATGAGTCAAGTATCGGCTCAGGGGTTCTTAATGCTTCT 1313
QY 249 CTGGCTCCAGTCTGGCAACAGGCTCTCTGACCAATCTTGGGCTCCAGGCTGAGGACG 308
DB 1312 CTGGCTCCAGTCTGGCAACAGGCTCTCTGACCAATCTTGGGCTCCAGGCTGAGGACG 1253
QY 309 AGGCTATTATTTACTGCTCATATATACAGAAAGTACTCT-----CATGCTTGGGA 362
DB 1252 AGGCTATTATTTACTGCTCATATATACAGAAAGTACTCTGATTTGTTCTTGGAA 1193
QY 363 CTGGGACCAAGTCAACGCTCTAGGTCAGCCCAAGGCAACCCCACTGTCACTCTGTTCC 422
DB 1192 CTGGGACCAAGTCAACGCTCTAGGTCAGCCCAAGGCTGAGCCCTCGGTCACTCTGTTCC 1133
QY 423 CGCCTCTCTGAGAGAGTCCCAAGCAAGGCAACACTAGTGTCTATCAGTACT 482
DB 1132 CGCCTCTCTGAGAGAGTCCCAAGCAAGGCAACACTAGTGTCTATTAAGTACT 1073
QY 483 TCTACCCGAGCTGTGACAGTGGCTGGAAGGAGATGGGAGCCGCTCAAGGCGGAG 542
DB 1072 TCTACCCGAGCGGTGACAGTGGCTGGAAGGAGATGGGAGCCGCTCAAGGCGGAG 1013
QY 543 TGGAGACCAACCAACCTCTCAACAGAGCAACCAAGTACGCGGCAAGCTACTGA 602
DB 1012 TGGAGACCAACCAACCTCTCAACAGAGCAACCAAGTACGCGGCAAGCTACTGA 953
QY 603 GCTGAGGCGCGGAGAGTGAAGTCCCAAGAGTCACTGTCAGAGTCAAGCATGAAG 662
DB 952 GCTGAGGCGCGGAGAGTGAAGTCCCAAGAGTCACTGTCAGAGTCAAGCATGAAG 893
QY 663 GAGACACCGTGAATGAAGACAGTGTCC 691
DB 892 GAGACACCGTGAATGAAGACAGTGTCC 865

RESULT 13
AADS9189/C
ID AADS9189 standard; cDNA; 2667 BP.
XX
AC AADS9189;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human rearranged Ig lambda chain cDNA.
XX
KW Colon cancer; gene therapy; human; Ig; immunoglobulin; ss.
XX
OS Homo sapiens.
XX
PN US2003073.05-A1.
XX
PD 17-APR-2003.
XX
PE 29-MAY-2002; 2002US-00158646.
XX
PR 31-MAY-2001; 2001US-0295239P.
XX
PA (LASEK/) LASEK A K W.
XX
PA (SORN/) SORNASSE T.
XX
PI Lasek AKW, Sornasse T;
XX
DR WPI; 2003-605964/57.
XX
PT Novel combination of cDNAs which are differentially expressed in colon
cancer; useful for detecting differential expression of one or more cDNAs

```

```

PT in a sample containing nucleic acid samples.
XX
XX Claim 1, Page 84-85; 89pp; English.
XX
XX The present invention relates to combination of cDNAs which are
CC differentially expressed in colon cancer. The invention is useful for
CC producing and purifying antibody, utilized as markers for treatment
CC efficacy against colon cancer. The invention is also useful for gene
CC therapy. The present sequence is human rearranged Ig lambda chain cDNA
XX
SQ Sequence 2667 BP; 602 A; 804 C; 721 G; 540 T; 0 U; 0 Other;
Query Match 39.8%; Score 581.2; DB 9; Length 2667;
Best Local Similarity 90.2%; Pred. No. 8.8e-127;
Matches 644; Conservative 0; Mismatches 66; Indels 2; Gaps 2;
QY 1 GGCCTGGGCTGCTATTTCTTCTACCCCTCTCACTGAGGCAACAGGCTCTGGGCCAGTC 60
DB 2633 GGCCTGGGCTGCTATTTCTTCTACCCCTCTCACTGAGGCAACAGGCTCTGGGCCAGTC 2574
QY 61 TGGCCTGACTAGTCTGCTCCGTGTCTGGTCTCTGAGAGTCAAGTCAACATCTCTG 120
DB 2573 TGGCCTGACTAGTCTGCTCCGTGTCTGGTCTCTGAGAGTCAAGTCAACATCTCTG 2514
QY 121 CACTGGAACCAAGCACTGAGTGTGTTATTAATATGCTCTCTGTAACCAACAGACC 180
DB 2513 CACTGGAACCAAGCACTGAGTGTGTTATTAATATGCTCTCTGTAACCAACAGACC 2454
QY 181 AGGCAAAAGCCCCAACTCATTTATGAGTCAATGAGGCTTCAAGGCTTCTTA 240
DB 2451 AGGCAAAAGCCCCAACTCATTTATGAGTCAATGAGGCTTCAAGGCTTCTTA 2394
QY 241 TCGCTTCTCTGCTCAAGTCTGCAACAGGCGCTCTGACATCTCT-GGGCTCCAGG 299
DB 2393 TCGCTTCTCTGCTCAAGTCTGCAACAGGCGCTCTGACATCTCTGAGGCTCCAGT 2334
QY 300 CTGAGAGCAGAGCTGATTTACTGCTGCTCATATACAGAAAGTACTTCTATGCTTTCG 359
DB 2333 CTGAGAGCAGAGCTGATTTACTGCTGCTCATATACAGAAAGTACTTCTATGCTTTCG 2274
QY 360 GAACCTGGGACCAAGGTCAACGCTCTAGGTCAGGCCCAAGGCCAACCCCACTGCTGT 419
DB 2273 GCGAGGAGCCAAAGTTGACCGTCTTAGGTCAGGCCCAAGGCCAACCCCACTGCTGT 2214
QY 420 TCCGCGCTCTCTGAGAGTCTCAAGCCCAAGGCCCACTAGTGTCTGATCACTG 479
DB 2213 TCCGCGCTCTCTGAGAGTCTCAAGCCCAAGGCCCACTAGTGTCTGATCACTG 2154
QY 480 ACTTTCTACCCGGGAGCTGTACAGTGTGCTGGAAGGCAAGTGGCAGGCCCTCAAGGCGG 539
DB 2153 ACTTTCTACCCGGGAGCTGTACAGTGTGCTGGAAGGCAAGTGGCAGGCCCTCAAGGCGG 2094
QY 540 GAGTGAAGACCAACCAACCTCTCAACAGAGCAACCAAGTACGCGGCCAGAGTACC 599
DB 2093 GAGTGAAGACCAACCAACCTCTCAACAGAGCAACCAAGTACGCGGCCAGAGTACC 2034
QY 600 TGAAGCTGAGCGCCGAGAGTGAAGTCCCAAGAGTCAAGTCAAGGCTCAAGCATG 659
DB 2033 TGAAGCTGAGCGCCGAGAGTGAAGTCCCAAGAGTCAAGTCAAGGCTCAAGCATG 1974
QY 660 AAGGAGAGCAGGTGATGAAGAGTGTCTCCCTACAGAAATGTTCTTAGGTT 713
DB 1973 AAGGAGAGCAGGTGATGAAGAGTGTCTCCCTACAGAAATGTTCTTAGGTT 1921

RESULT 14
AAS77073
ID AAS77073 standard; cDNA; 889 BP.
XX
AC AAS77073;
XX
DT 13-FEB-2002 (first entry)
XX

```





CC therapy. The present sequence is human Ig (hybridoma H210) anti-hepatitis  
CC A Ig lambda chain cDNA

xx Sequence 883 BP, 195 A; 283 C; 220 G; 185 T; 0 U; 0 Other;

Query Match 39.4%; Score 575.2; DB 9; Length 883;  
Best Local Similarity 87.9%; Pred. No. 1.8e-125;  
Matches 675; Conservative 0; Mismatches 83; Indels 10; Gaps 4;

```
QY 1 GGCCTGGGCTCTGCTATTCTCACCCTCTCTACTCAGGCGACAGGGTCTGGGCCCAATC 60
Db 41 GGCCTGGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 100
QY 61 TGCCCTGACTCAGTCTGCTCCGCTGCTGGGCTCTCTGACAGTCATCATCTCTG 120
Db 101 TGCCCTGACTCAGCTCTGCTCCGCTGCTGGGCTCTCTGACAGTCATCATCTCTG 160
QY 121 CACTGGAACCAAGTCACGTTGGTGTATTAATACTATGTCCTGTGTAACAACAACC 180
Db 161 CACTGGAAGTGGCGGTGACGTTGGTGTATTAATACTATGTCCTGTGTAACAACAACC 220
QY 181 AGGCAAGGCCCCCAAACTCATCATTTATTAAGTCAAGTATCGGCGCTCAGGGGTTCTAA 240
Db 221 AGGCAAGGCCCCCAAACTCATTTATTAATGATGATCAATATGGGCTCTCAGGGGTTCTAA 280
QY 241 TCGCTTCTGTGCTCAAGTCTGGCAACGCGCTCCCTGACATCTGTGGCTCCAGGC 300
Db 281 TCGCTTCTGTGCTCAAGTCTGGCAACGCGCTCCCTGACATCTGTGGCTCCAGGC 340
QY 301 TGAGGACGAGGCTATTTACTGCTGCTATATACA--AGAAGTACTTCTCATGTCTT 357
Db 341 TGAGGATAGGCTATTTACTGCTGCTATTTGAGTATGACATCTCTGTGTATT 400
QY 358 CGGAAGTGGGACCAAGTCAAGTCTAGGTCAGGCCCAAGGCCAACCCACTGTCACCTT 417
Db 401 CGGCGAGGAGACCAAGTCAAGTCTAGGTCAGGCCCAAGGCCAACCCACTGTCACCTT 460
QY 418 GTTCCCGCCTCTCTCTAGAGGCTCCAAAGCCCAAGGCCCACTAGTGTGTGATCAG 477
Db 461 GTTCCCGCCTCTCTCTAGAGGCTCCAAAGCCCAAGGCCCACTAGTGTGTGATCAG 520
QY 478 TGACTTTTACCGGGAGGCTGTGACAGTGGCTGTGAGGAGGAGGAGGCCCGTCAAGGC 537
Db 521 TGACTTTTACCGGGAGGCTGTGACAGTGGCTGTGAGGAGGAGGAGGCCCGTCAAGGC 580
QY 538 GGAAGTGAAGACCAACCAACCCCTCCAAACAGAGCAACACAAGTACCGGCGCAGAGCTA 597
Db 581 GGAAGTGAAGACCAACCAACCCCTCCAAACAGAGCAACACAAGTACCGGCGCAGAGCTA 640
QY 598 CTTGAGCTTGAAGCCCGAGCAGTGAAGTCCCAAGAACTTACAGTCTGACAGGCA 657
Db 641 TCTGAGCTTGAAGCCCGAGCAGTGAAGTCCCAAGAACTTACAGTCTGACAGGCA 700
QY 658 TGAAGGAGCAGCGTGAAGTGAAGCAGTGGTCCCTACAGAAATGTTCTATAGTTCTCT 717
Db 701 TGAAGGAGCAGCGTGAAGTGAAGCAGTGGTCCCTACAGAAATGTTCTATAGTTCTCT 755
QY 718 AACACGCTCAGCCCCCACTACCAATGCGAGACTGACACGCTGACGGG 765
Db 756 --AACCTCAGCCCCCACTACCAATGCGAGACTGACAGATCCACAG 801
```

Search completed: April 26, 2004, 16:12:07  
Job time : 560.73 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: April 26, 2004, 18:46:58 ; Search time 608.283 Seconds

(Without alignments)  
10821.600 Million cell updates/sec

Title: US-10-076-747-8

Perfect score: 1460  
Sequence: 1 ggcctggcctgtatcc.....ggcagcagcagcagcagct 1460

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 2907579 seqs, 225413464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCR\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCRUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1459	99.9	1460	US-10-076-747-8
2	809.6	55.5	957	US-10-076-747-7
3	599	41.0	14597	US-10-198-846-13540
4	597.4	40.9	1480	US-09-981-353-146
5	588.6	40.3	1640	US-10-198-846-13206
6	581.2	39.8	2667	US-10-158-846-76
7	575.6	39.4	1597	US-10-198-846-13540
8	575.2	39.4	883	US-10-158-846-73
9	565.4	38.7	729	US-10-461-148-12
10	529.4	36.3	726	US-10-225-108A-13
11	511.8	35.1	2112	US-10-001-857-108
12	505.8	34.6	2635	US-10-198-846-13529
13	498.2	34.1	819	US-10-342-887-1913
14	498.2	34.1	819	US-10-172-118-1913

15	497.2	34.1	605	US-03-878-134-264	Sequence 264, App
16	488.6	33.5	9472	US-03-837-308-194	Sequence 194, App
17	488.6	33.5	9532	US-10-045-674-451	Sequence 451, App
18	480.6	32.9	543	US-09-736-457-970	Sequence 970, App
19	480.6	32.9	543	US-09-902-941-970	Sequence 970, App
20	480.6	32.9	543	US-09-849-626-970	Sequence 970, App
21	480.6	32.9	543	US-10-283-017-970	Sequence 970, App
22	480.6	32.9	543	US-10-017-754-970	Sequence 970, App
23	480.6	32.9	543	US-10-113-872-970	Sequence 970, App
24	477.2	32.4	1640	US-10-198-846-13206	Sequence 13206, A
25	473.4	32.4	930	US-10-291-265-107	Sequence 107, App
26	473	32.4	848	US-10-158-646-70	Sequence 70, App
27	468.2	32.1	915	US-09-954-456-788	Sequence 788, App
28	468.2	32.1	915	US-09-880-107-3743	Sequence 3743, App
29	468.2	32.1	915	US-09-960-706-1069	Sequence 1069, App
30	468.2	32.1	915	US-09-873-319-710	Sequence 710, App
31	468.2	32.1	915	US-10-342-887-1912	Sequence 1912, App
32	468.2	32.1	915	US-10-172-118-1912	Sequence 1912, App
33	468.2	32.1	964	US-10-363-616-87	Sequence 87, App
34	462.8	31.7	888	US-10-158-646-71	Sequence 71, App
35	460	31.5	1590	US-10-198-846-12769	Sequence 12769, A
36	458	31.4	768	US-09-747-669-4	Sequence 4, App
37	458	31.4	768	US-09-747-669-5	Sequence 5, App
38	458	31.4	768	US-10-290-703-4	Sequence 4, App
39	458	31.4	768	US-10-290-703-5	Sequence 5, App
40	454.8	31.2	711	US-09-948-429B-9	Sequence 9, App
41	454.8	31.2	711	US-09-758-113-9	Sequence 9, App
42	454.8	31.2	711	US-10-124-905-9	Sequence 9, App
43	454.8	31.2	711	US-10-073-138-5	Sequence 5, App
44	454.8	31.2	711	US-10-124-807-9	Sequence 9, App
45	454.8	31.2	711	US-10-291-532-9	Sequence 9, App

#### ALIGNMENTS

RESULT 1  
US-10-076-747-8  
Sequence 8, Application US/10076747  
Publication No. US20030180726A1

GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Hu, Ping  
APPLICANT: Recipon, Hervé  
APPLICANT: Karra, Kalpana  
APPLICANT: Caffrey, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua

TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and  
FILE REFERENCE: DEX-0315  
CURRENT APPLICATION NUMBER: US/10/076,747  
CURRENT FILING DATE: 2002-02-13  
PRIOR APPLICATION NUMBER: 60/268,290  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 60/268,834  
PRIOR FILING DATE: 2001-02-15  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: PatentIn version 3.1

SEQ ID NO 8  
LENGTH: 1460  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1022)..(1022)  
OTHER INFORMATION: a, c, g or t

US-10-076-747-8

Query Match 99.9%; Score 1459; DB 15; Length 1460;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1460; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTGAGCTGTATTCTCACCCCTCTCACTCAGGAGCAAGGGGCTCTGGGCCAGTGC 60  
 Db 1 GGCCTGAGCTGTATTCTCACCCCTCTCACTCAGGAGCAAGGGGCTCTGGGCCAGTGC 60  
 QY 61 TCCCTGACTAGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Db 61 TCCCTGACTAGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 QY 121 CACTGGAACAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 Db 121 CACTGGAACAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 QY 181 AGGCAAAAGCCCCCAAACTCATATTATAGAGTCAAGTCAAGGCTCAAGGCTTCTTCA 240  
 Db 181 AGGCAAAAGCCCCCAAACTCATATTATAGAGTCAAGTCAAGGCTCAAGGCTTCTTCA 240  
 QY 241 TGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 Db 241 TGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 QY 301 TGAGACAGAGCTGATTTATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
 Db 301 TGAGACAGAGCTGATTTATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
 QY 361 AACTGGAGCAAGGTCAACCGTCTAGGTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 420  
 Db 361 AACTGGAGCAAGGTCAACCGTCTAGGTCAAGGCTCAAGGCTCAAGGCTCAAGGCTCA 420  
 QY 421 CCGGCTCTCTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 Db 421 CCGGCTCTCTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 QY 481 CTTCTACCGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 Db 481 CTTCTACCGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 QY 541 AGTGAAGACCAACAACTCTCAAAAGAGCAACAACTGACGAGTCAAGTCAAGTCA 600  
 Db 541 AGTGAAGACCAACAACTCTCAAAAGAGCAACAACTGACGAGTCAAGTCAAGTCA 600  
 QY 601 GAGCTGAGCGCCGAGAGTGAAGTCCACAGAAAGTCAAGTCAAGTCAAGTCAAGTCA 660  
 Db 601 GAGCTGAGCGCCGAGAGTGAAGTCCACAGAAAGTCAAGTCAAGTCAAGTCAAGTCA 660  
 QY 661 AGGAGCAACCGTGTAGTGAAGACAGTGTCTCTCAAGAAATGTTCTATAGTCTCTAAC 720  
 Db 661 AGGAGCAACCGTGTAGTGAAGACAGTGTCTCTCAAGAAATGTTCTATAGTCTCTAAC 720  
 QY 721 AGGCTGAGCGCCGAGAGTGAAGTCCACAGAAAGTCAAGTCAAGTCAAGTCAAGTCA 780  
 Db 721 AGGCTGAGCGCCGAGAGTGAAGTCCACAGAAAGTCAAGTCAAGTCAAGTCAAGTCA 780  
 QY 781 GGTCTCTCATCCACACACACAGAGGAGTCAAGGCTTATCATGCGCATGTGAC 840  
 Db 781 GGTCTCTCATCCACACACAGAGGAGTCAAGGCTTATCATGCGCATGTGAC 840  
 QY 841 GTCTATTAAGAGCGCTTACAAATTAAGATTTGTGCACTTGGTGAATGCTCTTACATCA 900  
 Db 841 GTCTATTAAGAGCGCTTACAAATTAAGATTTGTGCACTTGGTGAATGCTCTTACATCA 900  
 QY 901 TAAACAAGACATCTCACTCACTACACAGAGTCAAGGCTTATCATGCGCATGTGAC 960  
 Db 901 TAAACAAGACATCTCACTCACTACACAGAGTCAAGGCTTATCATGCGCATGTGAC 960  
 QY 961 GGCACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
 Db 961 GGCACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
 QY 1021 AACTATGAGTGAATTTGGCCGAGATCAAGACATGATTTTCCCGGAGTCAAGTCAAG 1080  
 Db 1021 AACTATGAGTGAATTTGGCCGAGATCAAGACATGATTTTCCCGGAGTCAAGTCAAG 1080  
 QY 1081 AATATGCGGTACGCGGACATCAATTCGCCCGGAGTCAAGACATTTAAAAAATGTGACGC 1140

Db 1081 AATATGCGGTACGCGGACATCAATTTGCCCGGAGTCAAGACATTTAAAAAATGTGACGC 1140  
 QY 1141 GCAGCGCGCGGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1200  
 Db 1141 GCAGCGCGCGGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1200  
 QY 1201 GGCAGCGCGGAGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1260  
 Db 1201 GGCAGCGCGGAGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1260  
 QY 1261 CGGAGCGCGGAGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1320  
 Db 1261 CGGAGCGCGGAGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1320  
 QY 1321 GCAGCGCGCGGAGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1380  
 Db 1321 GCAGCGCGCGGAGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1380  
 QY 1381 CAGCGCGCGGAGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1440  
 Db 1381 CAGCGCGCGGAGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 1440  
 QY 1441 GGCAGCGAGAGCGAGAGCT 1460  
 Db 1441 GGCAGCGAGAGCGAGAGCT 1460

## RESULT 2

US-10-076-747-7

Sequence 7, Application US/10076747

Publication No. US20030180726A1

GENERAL INFORMATION:

APPLICANT: Salceda, Susana

APPLICANT: Macina, Roberto

APPLICANT: Hu, Ping

APPLICANT: Recipon, Hervé

APPLICANT: Kariya, Kaipana

APPLICANT: Cafferykey, Robert

APPLICANT: Sun, Yongming

APPLICANT: Liu, Chenghua

TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and

FILE REFERENCE: DBX-0315

CURRENT APPLICATION NUMBER: US/10/076,747

CURRENT FILING DATE: 2002-02-13

PRIOR APPLICATION NUMBER: 60/268,290

PRIOR FILING DATE: 2001-02-13

PRIOR APPLICATION NUMBER: 60/268,834

PRIOR FILING DATE: 2001-02-15

NUMBER OF SEQ ID NOS: 129

SOFTWARE: PatentIn version 3.1

SEQ ID NO 7

LENGTH: 957

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: misc.feature

LOCATION: (519)..(519)

OTHER INFORMATION: a, c, g or t

US-10-076-747-7

Query Match 55.5%; Score 809.6; DB 15; Length 957;

Best Local Similarity 96.5%; Pred. No. 2.2e-230;

Matches 864; Conservative 0; Mismatches 19; Indels 12; Gaps 3;

QY 578 AAGTACGCGCGGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 633

Db 63 AAGTACGCGCGGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 633

QY 634 AAGTACGCGCGGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 685

Db 685 AAGTACGCGCGGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 685

QY 123 AAGTACGCGCGGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 182

Db 182 AAGTACGCGCGGAGAGAGAGCTTTTGAAGATGCCAACAAGAGGCGCGAGACCGGCA 182

```

QY 686 GGTCCCTACAGAAATGTTCTATAGTTCCTAACAACGCTCAGCCCCCACTACCAATGAC 745
Db 183 GGTCCCTACAGAAATGTTCTATAGTTCCTAACAACGCTCAGCCCCCACTACCAATGAC 242
QY 746 GAGACTAGCAGCTGCGAGGATCCCAAGGAGAGGGGTCTCTCCATCCACACCAAG 805
Db 243 GAGACTAGCAGCTGCGAGGATCCCAAGGAGAGGGGTCTCTCCATCCACACCAAG 302
QY 806 GGGGAGTCAAGGCTTATATGCGGCAATGCAACGTCATGTAAGGCGCTCAATAA 865
Db 303 GGGGAGTCAAGGCTTATATGCGGCAATGCAACGTCATGTAAGGCGCTCAATAA 362
QY 866 GATATTCGACCTTGTAATGCTCTACATACATAACAAGACATCTCACTACATACA 925
Db 363 GATATTCGACCTTGTAATGCTCTACATACATAACAAGACATCTCACTACATACA 422
QY 926 CGGAGTGAATGATCACCCTGCTGCGGCGAATGCCACTTGCTCTGCTGCTCAAT 985
Db 423 CGGAGTGAATGATCACCCTGCTGCGGCGAATGCCACTTGCTCTGCTGCTCAAT 482
QY 986 GATGACTCATGTATACACACCGGTGTGCGGCGACANCTATGATGATGATTTGCGCGGG 1045
Db 483 GATGACTCATGTATACACACCGGTGTGCGGCGACANCTATGATGATGATTTGCGCGGG 542
QY 1046 ATCAACAGACCATGATTTCCCGCTGCTGATGATGATGATGATGATGATGATGAT 1105
Db 543 ATCAACAGACCATGATTTCCCGCTGCTGATGATGATGATGATGATGATGATGATGAT 602
QY 1106 CGCCCGGAGCTATAGACCTTAAAAAGTTGACCGCGGAGCGGCGGAGGAGGAGGAGG 1165
Db 603 CGCCCGGAGCTATAGACCTTAAAAAGTTGACCGCGGAGCGGCGGAGGAGGAGGAGG 662
QY 1166 ACCGAAATGCCAAACACAGAGGCGCGACACCGGAGGCGACCCCGGCGGAGGAGGCGCA 1225
Db 663 ACCGAAATGCCAAACACAGAGGCGCGACACCGGAGGCGACCCCGGCGGAGGAGGCGCA 722
QY 1226 CTGCGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1285
Db 723 CTGCGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 782
QY 1286 TGAACAGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1345
Db 783 TGAACAGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 842
QY 1346 AGCCACACGCGCTTATGCGGCGACGCGCGGCGACGCGCGGCGAGGAGGCGGAGGAG 1405
Db 843 AGCCACACGCGCTTATGCGGCGACGCGCGGCGACGCGCGGCGAGGAGGCGGAGGAG 902
QY 1406 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1460
Db 903 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 957

```

```

RESULT 3
US-10-198-846-13540/c
; Sequence 13540, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198, 846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306, 220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13540

```

```

; LENGTH: 1597
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-13540
Query Match 41.0%; Score 599; DB 15; Length 1597;
Best Local Similarity 92.6%; Pred. No. 1,3e-167;
Matches 640; Conservative 0; Mismatches 50; Indels 1; Gaps 1;
QY 1 GGCTTGGCTCTGCTATTCCTACCTCTCTCACTCAGGCGCAGGGTCTTGGCCCACTC 60
Db 1528 GGCTTGGCTCTGCTATTCCTACCTCTCTCACTCAGGCGCAGGGTCTTGGCCCACTC 1469
QY 61 TGCCCTGACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 1468 TGCCCTGACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1409
QY 121 CACTGGAACAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 1408 CACTGGAACCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1349
QY 181 AGGCAAAAGCCCCCAATCATCATTTATGAGTGAATATGAGGCTTCAAGGGGTTTCTAA 240
Db 1348 AGGCAAAAGCCCCCAATCATCATTTATGAGTGAATATGAGGCTTCAAGGGGTTTCTAA 1289
QY 241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 1288 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1229
QY 301 TGAGGAGAGGCTGATTTATCTGCTGCTCATATACAGAGTATCTTCAATGCTTCTGCT 360
Db 1228 TGAGGAGAGGCTGATTTATCTGCTGCTCATATACAGAGTATCTTCAATGCTTCTGCT 1169
QY 361 AACTGGAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 1168 AACTGGAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1109
QY 421 CCGGCTCTCTTGAAGGCTTCAAGCCCAAGGCGCACTATGCTGCTGATGATGATGAT 480
Db 1108 CCGGCTCTCTTGAAGGCTTCAAGCCCAAGGCGCACTATGCTGCTGATGATGATGAT 1049
QY 481 CTTTACCCGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 1048 CTTTACCCGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
QY 541 AGTGAAGACCAACCAACCTTCAACAGAGCAACCAAGTACGCGGCGAGAGTACTCT 600
Db 988 AGTGAAGACCAACCAACCTTCAACAGAGCAACCAAGTACGCGGCGAGAGTACTCT 929
QY 601 GAGCTGACGCGGCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 928 GAGCTGACGCGGCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 869
QY 661 AGGAGACACCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 691
Db 868 AGGAGACACCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839

```

```

RESULT 4
US-09-981-353-146/c
; Sequence 146, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 146
; LENGTH: 1480

```

TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Inocyte ID No. US20020160382A1 1329606.3  
LOCATION: 134, 198, 206  
OTHER INFORMATION: a, t, c, g, or other  
US-09-981-353-146

Query Match 40.9%; Score 597.4; DB 9; Length 1480;  
Best Local Similarity 93.1%; Pred. No. 3.9e-167;  
Matches 649; Conservative 0; Mismatches 41; Indels 7; Gaps 2;

QY 1 GGCCTGGGCTGCTGCTATTCCTCACTCCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 60  
DB 1441 GGCCTGGGCTGCTGCTATTCCTCACTCCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 1382  
QY 61 TGCCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 1381 TGCCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1322  
QY 121 CACTGGAACGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
DB 1321 CACTGGAACGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1262  
QY 181 AGGCAAGGCCCCCAAACTCATTTATGAGTCAATGAGGCTCAGGCTCAGGCTTTCTAA 240  
DB 1261 AGGCAAGGCCCCCAAACTCATTTATGAGTCAATGAGGCTCAGGCTCAGGCTTTCTAA 1202  
QY 241 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 1201 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1142  
QY 301 TGAAGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 354  
DB 1141 TGAAGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082  
QY 355 CTTCGGAACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414  
DB 1081 CTTCGGAACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022  
QY 415 TCTGTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 474  
DB 1021 TCTGTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962  
QY 475 CAGTACTTCTACCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 534  
DB 961 CAGTACTTCTACCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902  
QY 535 GCGGGAAGTGAAGCACAACCTCCAAACAGAGCAACAAGTACCGGCTCAGAG 594  
DB 901 GCGGGAAGTGAAGCACAACCTCCAAACAGAGCAACAAGTACCGGCTCAGAG 842  
QY 595 CTACTCTGAGGCTGAGCGCCGAGCAGTGAAGTCCCAAGAGCTACAGCTCCAGGTAC 654  
DB 841 CTACTCTGAGGCTGAGCGCCGAGCAGTGAAGTCCCAAGAGCTACAGCTCCAGGTAC 782  
QY 655 GCATGAGGAGGACACCGTGAATGAAGACAGTGTCCC 691  
DB 781 GCATGAGGAGGACACCGTGAATGAAGACAGTGTCCC 746

RESULT 5  
US-10-198-846-13206/c  
Sequence 13206, Application US/10198846  
Publication No. US2003009974A1  
GENERAL INFORMATION:  
APPLICANT: Lillie, James  
APPLICANT: Xu, Yongyao  
APPLICANT: Wang, Youzhen  
APPLICANT: Steinmann, Kathleen  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
TITLE OF INVENTION: THERAPY OF BREAST CANCER  
FILE REFERENCE: MRI-049  
CURRENT APPLICATION NUMBER: US/10/198,846  
CURRENT FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/306,220  
PRIOR FILING DATE: 2001-07-18  
NUMBER OF SEQ ID NOS: 14084  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13206  
LENGTH: 1640  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-198-846-13206

Query Match 40.3%; Score 588.6; DB 15; Length 1640;  
Best Local Similarity 89.4%; Pred. No. 1.7e-164;  
Matches 657; Conservative 0; Mismatches 74; Indels 4; Gaps 2;

QY 1 GGCCTGGGCTGCTGCTATTCCTCACTCCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 60  
DB 1567 GGCCTGGGCTGCTGCTATTCCTCACTCCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 1508  
QY 61 TGCCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 1507 TGCCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1448  
QY 121 CACTGGAACGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
DB 1447 CACTGGAACGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1386  
QY 181 AGGCAAGGCCCCCAAACTCATTTATGAGTCAATGAGGCTCAGGCTCAGGCTTTCTAA 240  
DB 1387 AGGCAAGGCCCCCAAACTCATTTATGAGTCAATGAGGCTCAGGCTCAGGCTTTCTAA 1328  
QY 241 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 1327 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1268  
QY 301 TGAAGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
DB 1267 TGAAGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1208  
QY 361 AACTGGAGCAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
DB 1207 CCGAGGAGCAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1148  
QY 421 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
DB 1147 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1088  
QY 481 CTTCCTACCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
DB 1087 CTTCCTACCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1028  
QY 541 AGTGAAGACCAACCAACCTCCAAACAGAGCAACAAGTACCGGCTCAGAGTACT 600  
DB 1027 AGTGAAGACCAACCAACCTCCAAACAGAGCAACAAGTACCGGCTCAGAGTACT 968  
QY 601 GAGCTGAGCGCCGAGCAGTGAAGTCCCAAGAGCTACAGCTCCAGGTACGATGA 660  
DB 967 GAGCTGAGCGCTGAGCAGTGAAGTCCCAAGAGCTACAGCTCCAGGTACGATGA 908  
QY 661 AGGAGACAGGCTGATGAAGACAGTGTCCCTACAGAAATTTCTATAGGTTCTCTAAC 720  
DB 907 AGGAGACAGGCTGATGAAGACAGTGTCCCTACAGAAATTTCTATAGGTTCTCTAAC 852  
QY 721 ACGCTCAGCCCCAC 735  
DB 851 CTCACCCCCCCCCAC 837

RESULT 6



```

Db      637 GCTACTAGAGCTGAGCTGAGCAGTGGAGTCCCAAGAGCTACAGCTGCCAGGTCA 696
QY      654 GCGATGAAGGAGAGCAGCCTGGATGAAGACAGTGGTCCCTACAGAAATGTTCTAGATT 713
Db      697 GCGATGAAGGAGAGCAGCCTGGAG-GAAGACAGTGGGCCC--TACAGATGTTCTAGATT 752
QY      714 CTCTACAGAGCTCAGCCCCCACTACCAATGGGAGAGTACAGACGCTGAGGG 765
Db      753 CTC--AACCTCACCCCCCACCAGGAGACTAGAGCTGAGATCCCAAG 801

RESULT 8
US-10-158-646-73
; Sequence 73, Application US/10158646
; Publication No. US20030073105A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Sornasse, Thierry
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0030-1 US
; CURRENT APPLICATION NUMBER: US/10/158,646
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/295,239
; PRIOR FILING DATE: 2001-05-31
; SOFTWARE: PERL Program
; SEQ ID NO 73
; LENGTH: 883
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030073105A1 1329913.2
US-10-158-646-73

Query Match      39.4%; Score 575.2; DB 15; Length 883;
Best Local Similarity 87.9%; Pred. No. 1.3e-160;
Matches 675; Conservative 0; Mismatches 83; Indels 10; Gaps 4;

QY      1 GGCCTGGGCTGTCTATTCTCTACCCCTCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 60
Db      41 GGCCTGGGCTGTCTCTCTCTCTCACTCTCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 100
QY      61 TGCCTGACTCAGTGTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 120
Db      101 TGCCTGACTCAGTGTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 160
QY      121 CACTGGAACAGCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db      161 CACTGGAAGTGGCGGTGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 220
QY      181 AGGCAAGCCCCCAACTCATCATTTATGAGTCACTGATGAGGCTCTGAGGGGTTCTAA 240
Db      221 AGGCAAGCCCCCAACTCATCATTTATGAGTCACTGATGAGGCTCTGAGGGGTTCTAA 280
QY      241 TGGCTTCTGGCTCAAGTCTGCAAGAGGAGGCTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db      281 TGGCTTCTGGCTCAAGTCTGCAAGAGGAGGCTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 340
QY      301 TGAGAGCAGAGGCTGATTAATTAATGCTGCTCATATACA--AGAAGTCTTCTCATGCTT 357
Db      341 TGAGAGTGAAGGCTTATTAATTAATGCTGCTCATATTAAGTGAAGTGAAGTCTTCTGTTATT 400
QY      358 CCGAAGTGGAGACAGAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 417
Db      401 CCGAGAGGAGACAGAGTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 460
QY      418 GTTCCGCGCTCTCTGAGAGCTCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 477
Db      461 GTTCCGCGCTCTCTGAGAGCTTCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 520
QY      478 TGACTTTACCCGGAGGCTGTGACAGTGGCTGGAAGGAGAGTGGACAGCCCGTCAAGGC 537

```

```

Db      521 TGACTTTACCCGGAGGAGCCTGTGACAGTGGCTTGAAGGAGATAGACAGCCCGCTGAGGC 580
QY      538 GGGAGTGGAGACCAACCAACCTCTCAACAGAGCAACAAAGTACGCGGCCAGAGCTA 597
Db      581 GGGAGTGGAGACCAACCAACCTCTCAACAGAGCAACAAAGTACGCGGCCAGAGCTA 640
QY      598 CCGAGCTGAGCCCGAGCAGTGAAGTCCAGAGAAAGCTACAGCTGCAGGTCAAGCA 657
Db      641 TCTGAGCTGAGCCTGAGCAGTGAAGTCCAGAGAAAGCTACAGCTGCAGGTCAAGCA 700
QY      658 TGAAGGAGACCGTGGATGAAGACAGTGTCCCTCAAGAAATGTTTAGTTCCT 717
Db      701 TGAAGGAGACCGTGGAG-GAAGACAGTGGCCCC--TACAGAAATGTTTAGTTCCT- 755
QY      718 AACAGCTCAGCCCGCAGTACCAATGGGAGAGCTAGCAGCGTCAAGG 765
Db      756 --AACCTCACCCCCCACCAGGAGACTAGAGCTGAGAGATCCAGG 801

RESULT 9
US-10-461-148-12
; Sequence 12, Application US/10461148
; Publication No. US20040013672A1
; GENERAL INFORMATION:
; APPLICANT: Dietzschold, Bernhard
; APPLICANT: Hooper, Douglas C.
; TITLE OF INVENTION: RECOMBINANT ANTIBODIES AND COMPOSITIONS
; FILE REFERENCE: 8321-110C11-185685
; CURRENT APPLICATION NUMBER: US/10/461,148
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 10/225,108
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/314,023
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/848,832
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/204,518
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Human
US-10-461-148-12

Query Match      38.7%; Score 565.4; DB 16; Length 729;
Best Local Similarity 90.2%; Pred. No. 1e-157;
Matches 629; Conservative 0; Mismatches 61; Indels 7; Gaps 2;

QY      1 GGCCTGGGCTGTCTATTCTCTACCCCTCTCTCACTCAGGAGCAGAGGCTCTGGGCCAGTC 60
Db      18 GGCCTGGGCTGTCTCTCTCTCTCACTCTCTCACTCAGGAGCAGAGATCTGGGCTCACTG 77
QY      61 TGCCTGACTCAGTGTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTG 120
Db      78 TGCCTGACTCAGTGTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTG 137
QY      121 CACTGGAACAGCAGTCAAGTGTGCTTATACTAGTCTTCTGTAACAAAGCAGACCC 180
Db      138 CACTGGAACAGCAGTATGATGCTTATTAATTTCTCTCTGTAACAAAGCAGACCC 197
QY      181 AGGCAAGCCCCCAACTCATTTATGAGTCACTGATGAGGCTCTGAGGGGTTCTAA 240
Db      198 AGGCAAGCCCCCAACTCATTTATGATGCTTATTAAGTCACTGAGGCTCTGAGGGGTTCTGA 257
QY      241 TGCCTTCTGCTCAAGTCTGCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 300
Db      258 TGCCTTCTGCTCAAGTCTGCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 317
QY      301 TGAGAGCAGAGGCTGATTAATTAATGCTGCTCATATACAG-----AAGTACTTCTCATAGT 354

```

```

Db      318 TGAGGATAGAGGCTGATTTATTAATCTGTGTCATATGACGCGCATACACCCCGGGCGTGT 377
Qy      355 CTTCGGAATCGGAGCAAGAGTCAACCGTCTTAGGTGAGCCCAAGGCCAACCCCACTGTAC 414
Db      378 TTTCGGCGGAGGAGCAAGAGTCAACCGTCTTAGGTGAGCCCAAGGCCAACCCCACTGTAC 437
Qy      415 TCTGTTCCCGCCCTCTCTCTGAGAGAGTCCAAAGCCCAAGGCCCACTAGTGTGTGAT 474
Db      438 TCTGTTCCCGCCCTCTCTCTGAGAGAGTCCAAAGCCCAAGGCCCACTAGTGTGTGAT 497
Qy      475 CAGTACTTCTACCCGGGAGCTGTGACAGTGGCCCTGGAAGGCAATGGACCCCGTCAA 534
Db      498 AAGTACTTCTACCCGGGAGCTGTGACAGTGGCCCTGGAAGGCAATGGACCCCGTCAA 557
Qy      535 GGGCGGAGTGAAGACCAACCAACCCCTCCAAACAGAGCAACAAAGTACGGGGCAGAG 594
Db      558 GGGCGGAGTGAAGACCAACCAACCCCTCCAAACAGAGCAACAAAGTACGGGGCAGAG 617
Qy      595 CTACTGTAGCTGAGCCCGGAGCAGTGAAGTCCCAAGAGCTACAGCTGCCAGGTAC 654
Db      618 CTACTGTAGCTGAGCCCGGAGCAGTGAAGTCCCAAGAGCTACAGCTGCCAGGTAC 677
Qy      655 GCATGAAGGAGCAACCGTGAAGTGAAGAGTGTGCTCC 691
Db      678 GCATGAAGGAGCAACCGTGAAGTGAAGAGTGTGCTCC 713

```

```

RESULT 10
US-10-225-108A-13
; Sequence 13, Application US/10225108A
; Publication No. US20030157112M1
; GENERAL INFORMATION:
; APPLICANT: HOOPER, Craig
; APPLICANT: DIETZSCHOLD, Bernhard
; TITLE OF INVENTION: Recombinant Antibodies, and Compositions
; FILE REFERENCE: 8321-110
; CURRENT APPLICATION NUMBER: US/10/225-108A
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US 09/848,832
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/204,518
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/314,023
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-108A-13

```

```

Query Match      36.3%; Score 529.4; DB 15; Length 726;
Best Local Similarity 89.8%; Pred. No. 5.3e-147;
Matches 626; Conservative 0; Mismatches 61; Indels 10; Gaps 5;

```

```

Qy      1 GGCCTGAGCTGCTGCTATTCCTCAACCTCCCTCACTCAGGAGCAAGGGCTCTGGGCCAGTGC 60
Db      18 GGCCTGAGCTGCTGCTATTCCTCAACCTCCCTCACTCAGGAGCAAGGGCTCTGGGCCAGTGC 77
Qy      61 TGCCCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db      78 TGCCCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 137
Qy      121 CACTGGAACCAAGAGTCAAGTGTGTGTATTAATATGTCTCTGCTGCTGCTGCTGCTGCTGCTG 180
Db      138 CACTGGAACCAAGAGTGTGTGTGTATTAATATGTCTCTGCTGCTGCTGCTGCTGCTGCTG 197
Qy      181 AGGCAAGAGCCCAACATCATTTATGAGTCAAGTATCGGCGCTCAGAGGGTTCTTA 240
Db      198 AGGCAAGAGCCCAACATCATTTATGAGTCAAGTATCGGCGCTCAGAGGGTTCTTA 257

```

```

Qy      241 TCGCTTCTGTGCTCCAGAGTCTGCAACAGGCGCTCCCTGACCATCTGTGGCTCCAGAGC 300
Db      258 TCGCTTCTGTGCTCCAGAGTCTGCAACAGGCGCTCCCTGACCATCTGTGGCTCCAGAGC 317
Qy      301 TGAAGACAGAGGCTGATTTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
Db      318 TGAAGACAGAGGCTGATTTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377
Qy      355 CTTCGGAATCGGAGCAAGAGTCAACCGTCTTAGGTGAGCCCAAGGCCAACCCCACTGTAC 414
Db      378 TTTCGGCGGAGGAGCAAGAGTCAACCGTCTTAGGTGAGCCCAAGGCCAACCCCACTGTAC 437
Qy      415 TCTGTTCCCGCCCTCTCTCTGAGAGAGTCCAAAGCCCAAGGCCCACTAGTGTGTGAT 474
Db      438 TCTGTTCCCGCCCTCTCTCTGAGAGAGTCCAAAGCCCAAGGCCCACTAGTGTGTGAT 497
Qy      475 CAGTACTTCTACCCGGGAGCTGTGACAGTGGCCCTGGAAGGCAATGGACCCCGTCAA 534
Db      498 AAGTACTTCTACCCGGGAGCTGTGACAGTGGCCCTGGAAGGCAATGGACCCCGTCAA 556
Qy      535 GGGCGGAGTGAAGACCAACCAACCCCTCCAAACAGAGCAACAAAGTACGGGGCAGAG 594
Db      557 GGGCGGAGTGAAGACCAACCAACCCCTCCAAACAGAGCAACAAAGTACGGGGCAGAG 615
Qy      595 CTACTGTAGCTGAGCCCGGAGCAGTGAAGTCCCAAGAGCTACAGCTGCCAGGTAC 654
Db      616 CTACTGTAGCTGAGCCCGGAGCAGTGAAGTCCCAAGAGCTACAGCTGCCAGGTAC 674
Qy      655 GCATGAAGGAGCAACCGTGAAGTGAAGAGTGTGCTCC 691
Db      675 GCATGAAGGAGCAACCGTGAAGTGAAGAGTGTGCTCC 710

```

```

RESULT 11
US-10-001-857-108
; Sequence 108, Application US/10001857
; Publication No. US20020183500A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Helve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenchua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro
; FILE REFERENCE: DEX-0273
; CURRENT APPLICATION NUMBER: US/10/001-857
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,054
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 108
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2005)..(2005)
; OTHER INFORMATION: a, c, g or t
US-10-001-857-108

```

```

Query Match      35.1%; Score 511.8; DB 14; Length 2112;
Best Local Similarity 87.2%; Pred. No. 1.3e-141;
Matches 585; Conservative 0; Mismatches 82; Indels 4; Gaps 2;

```

```

Qy      24 CCTCTCTACTCAGGAGCAAGGGTCTGGGCCCACTGTGCTGCTGCTGCTGCTGCTGCTGCTG 83
Db      779 CCTTGGCCCAAGACCTCTAGGGTCTGGGCCCACTGTGCTGCTGCTGCTGCTGCTGCTG 838
Qy      84 TGTCTGGGTCCTGGGAGAGTGCATCAACATCTCTGTGCTGGAACCAAGAGTCAAGTTC 143
Db      839 TGTCTGGGTCCTGGGAGAGTGCATCAACATCTCTGTGCTGGAACCAAGTTCACATTCG 898

```



QY	144	G---TGGTATATACATATGCTCCTGGTATACCAACAGACACCAGAGCAAAAGCCCCAAATCA	200
Db	899	GGGACGGTTATATACATATGACACTGGTATACACAGCTTCCAGAAACAGCCCCAAATCA	958
QY	201	TCATTTATGAGGTGCAATATCGACCCCTCAAGGGGTTTCTAATGCTTCTTGCTCCAACT	260
Db	959	TGATTTATGAGGTGCTGTAAGCAGACCCCTCAGGGGGTTTCTGATGCTTCTTGCTCCAACT	1018
QY	261	CTGGCAACACGGCCCTTCCTGACCAATCTCTGGGCTCCAGGCTGAGACAGAGCTGATTAT	320
Db	1019	CTGGCAACACGGCCCTTCCTGACCAATCTCTGGGCTCCAGGCTGAGACAGAGCTGATTAT	1078
QY	321	ACTGCTGCTCATATACAAAGATACCTCTCATGCTGTGGAACTGGACCAAGGTCACCG	380
Db	1079	ACTGCTGCTCATATGCAAGGACTCTCATCTTGAGTGTGGCGAGAGACCAAGCTGACCG	1138
QY	381	TCCTTAGTCAAGCCCAAGGCCAACCCCACTGTCACCTGTCTCCGCGCTCTCTGAGAGC	440
Db	1139	TCCTTAGTCAAGCCCAAGGCTGCGCCCTGGGTCACTGTGTCCGCGCTCTCTGAGAGC	1198
QY	441	TCCAAGCCCAACAGGCCACACTAGTGTGTCTGATCAGTGACTTCTAACCCGGAGCTGTGA	500
Db	1199	TTCAAGCCCAACAGGCCCACTAGTGTGTCTCATAGTGACTTCTAACCCGGAGCCGTGA	1258
QY	501	CAGTAGCCCTGGAAGGCGAGATGGCAACCCCGTCAAGCGCGAGTGGAGACCAACCAACCT	560
Db	1259	CAGTAGCCCTGGAAGGCGAGATGGCAACCCCGTCAAGCGCGAGTGGAGACCAACCAACCT	1318
QY	561	CCAAACAGAGCAACACAAATACGCGGCGACAGCTACCTGAGCTGAGCGCCGAGCAGT	620
Db	1319	CCAAACAAAGCAACACAAATACGCGGCGACAGTACTGAGCTGAGCGCTGAGCAGT	1378
QY	621	GGAACTCCCAAGAACTACAGCTGCCAGGTCAAGCATGAAGAGAGACACCTGTGATGAAG	680
Db	1379	GGAACTCCCAAGAACTACAGCTGCCAGGTCAAGCATGAAGAGAGACACCTGTGA- GAA	1437
QY	681	ACAGTGTGCC 691	
Db	1438	ACAGTGTGCC 1448	

RESULT 14  
US-10-172-118-19.3  
Sequence 1913, Application US/1012118  
Publication NO. US2003022437A1  
GENERAL INFORMATION:  
APPLICANT: Dai, Hongyue  
APPLICANT: He, Yudong  
APPLICANT: Linsley, Peter  
APPLICANT: Mao, Mao  
APPLICANT: Roberts, Chris  
APPLICANT: Van 't Veer, Laura  
APPLICANT: Van de Vijver, Marc

QY	1	GGGCTTGGGCTCGATATTCACCTCTCACTAGAGGACAGAGGTCTTGGGCCAGTC	60
Db	1	GGGCAAGCTTCCCTCTCTCTCTCACTCTCTCACTCACTGTGCAAGGCTCTGGGCCAGTC	60
QY	61	TGCGCTGACTCAGTGTGCTCCGTGTGTGGGTCTTCCTGGACAGTCGATCACCATCTCTG	120
Db	61	TGNGCTGACTCAGCACCTTCAGCTGTGGAGCCCGGGCAGAGGGTCATCATCTCTTG	120
QY	121	CAGTGGAAACACAGATCAGCTTGTGTGTATATCATGTCTCCGTGGTACCAACAGCACCC	180
Db	121	TTCTGGAAACCACTCCACATCGAGGGTAAATA--CTGTAACTGTATCCAGCAGCTCCC	177
QY	181	AGGCAAGGCCCCCAACTCATCATTTATGAGGTCAAGTATCGGCGCTCAGGGGTTCTTAA	240
Db	178	AGGAAGGCGCCCCCAACTCTCATCATATGATATATAGGGGCGCTCAGGGGTCCTGA	233
QY	241	TGCGTTCTTGTGCTCOAAGTCTGGCAACAGGCGCTCCCTGACCATCTCTGGGCTTCAGGC	300
Db	238	CCGATTTCTTGTGCTCOAAGTCTGGCACTCAGCTCCCTCGCCATCATGTGGGCTTCAGTC	297
QY	301	TGAGCAGAGGCTGATTTACTGTGCTCAT-----ATACAAAGATATTTCTCATGT	354
Db	298	TGAGGATGAGGCTGATTTACTGTGACGACATGGGATGACAGCTGAAATGTCGATATGT	353
QY	355	CTTCGAACTGGAGCAAGTCAACCGTCTTAGTCAAGCCCAAGGCCAACCCCACTGTCAAC	414
Db	358	CTTCGGAACTGGAGCAAGGTCAACCGTCTTAGTCAAGCCCAAGGCCAACCCCACTGTCAAC	417
QY	415	TCTGTTCCCGCCCTCTCTGTAGAGAGCTCCAGCCCAAGGCCCACTTAGTGTCTGAT	476
Db	418	TCTGTTCCCGCCCTCTCTGTAGAGAGCTCCAAAGCCCAAGGCCCACTTAGTGTCTGAT	477
QY	475	CAGTCACTTCTACCCGGAGCTGTGACAGTGTGCTGGAGGCAAGTGGACGCCCTGTCAA	534
Db	478	CAGTCACTTCTACCCGGAGCTGTGACAGTGTGCTGGAGGCAATGGCAGCCCGTCAA	537
QY	535	GGCGGGAGTGGAGACCAACCAACCTCTCCAAACAGACCAACCAAGTACGCGGCTCAGAG	594
Db	538	GGCGGGAGTGGAGACCAACCAACCTCTCCAAACAGACCAACCAAGTACCGGGCAGACAG	597
QY	595	CTTACCTGAGCTCGACGCCCGCAGCAGTGGAAATGCCACAGAAAGTACAGCTGCCAGGTCAAC	654
Db	598	CTTACTGAGCTCGAGGCCCGCAGCAGTGGAAATGCCACAGAAAGTACAGCTGCCAGGTCAAC	655
QY	655	GCATGAAAGGAGCAACCGTGAATGAAGACAGTGTCCC	691
Db	658	GCATGAAAGGAGCAACCGTGA--GAAGACAGTGTCCCC	693

RESULT 15  
US-09-878-134-264  
; Sequence 264, Application US/09878134

```
/ Publication No. US20020086303A1
/ GENERAL INFORMATION:
/ APPLICANT: Meagher, Madeleine Joy
/ APPLICANT: King, Gordon E.
/ APPLICANT: Xu, Jiangchun
/ APPLICANT: Secrist, Heather
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
/ FILE REFERENCE: 210121.532
/ CURRENT APPLICATION NUMBER: US/09/878,134
/ CURRENT FILING DATE: 2001-06-07
/ NUMBER OF SEQ ID NOS: 377
/ SOFTWARE: Corixa Invention Disclosure Database
/ SEQ ID NO 264
/ LENGTH: 605
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(605)
/ OTHER INFORMATION: n = A,T,C or G
US-09-878-134-264
```

```
Query Match 34.1%; Score 497.2; DB 13; Length 605;
Best Local Similarity 92.8%; Pred. No. 1.9e-137;
Matches 531; Conservative 0; Mismatches 40; Indels 1; Gaps 1;
```

```
QY 1 GGCCTGGGCTGCTATTCTCCACCCCTCCTCAGGGGACAGGGGCTCGGGCCAGTC 60
|
|
|
Db 35 GGCCTGGGCTGCTATTCTCCACCCCTCCTCAGGGGACAGGGGCTCGGGCCAGTC 94
|
|
|
QY 61 TGCCCTGACTAGTTCCTCCGCTCTGCTGCTCTGCTCTGCTGCTGCTGCTGCTGCT 120
|
|
|
Db 95 TGCCCTGACTAGTTCCTCCGCTCTGCTGCTCTGCTGCTCTGCTGCTGCTGCTGCTG 154
|
|
|
QY 121 CACTGGAACAGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 180
|
|
|
Db 155 CACTGGAACAGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 214
|
|
|
QY 181 AGGCAAGGCCCCCAACTCATCTATTATAGGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240
|
|
|
Db 215 AGGCAAGGCCCCCAACTCATCTATTATAGGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 274
|
|
|
QY 241 TGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
|
|
|
Db 275 TGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 334
|
|
|
QY 301 TGAGGACGAGGCTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 360
|
|
|
Db 335 TGAGGACGAGGCTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 394
|
|
|
QY 361 AACTGGACCAAGTTCACCGTCTAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 420
|
|
|
Db 395 CGAGGGGACCAAGTTCACCGTCTAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 454
|
|
|
QY 421 CCGGCTCTCTGAGGAGCTTCAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 480
|
|
|
Db 455 CCGGCTCTCTGAGGAGCTTCAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 514
|
|
|
QY 481 CTTCTACCGGAGGCTGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 540
|
|
|
Db 515 CTTCTACCGGAGGCTGAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 573
|
|
|
QY 541 AGTGAAGACCAACCAACCTCTCAACAGAGCA 572
|
|
|
Db 574 AGTGAAGACCAACCAACCTCTCAACAGAGCA 605
|
|
|
```

Search completed: April 26, 2004, 22:54:12  
Job time : 612.283 secs

LIBRARY: THYNOT10

CLONE: 2872705  
US-09-049-672A-23

US-09-049-672A-23

Query Match	41.6%;	Score 607;	DB 3;	Length 891;
Best Local Similarity	93.3%;	Pred. No. 6.8e-149;		
Matches 645;	Conservative	0;	Mismatches 45;	Indels 1;
			Gaps	1;

QY	GGCTCTGGGCTCTGTATTTCTCACTCCCTCTCACTCAGAGGACAGAGGCTCTGGAGCCAGTGC	60	
Db	36	GGCTCTGGGCTCTGTATTTCTCACTCCCTCTCACTCAGAGGACAGAGGCTCTGGAGCCAGTGC	95
QY	61	TGCGCTTACCTAGTCTGCGCTCGGTGTCTGGGCTCTCTCTGGACAGTGCATCACATCTCTTG	120
Db	96	TGCGCTTACCTAGCTGCGCTCGGTGTCTGGGCTCTCTCTGGACAGTGCATCACATCTCTTG	155
QY	121	CACATGAAACAGCAGTCAAGTTGGTGGTTATTAACATATGTCTCTGGTACCAAGCAACCC	180
Db	156	CACATGAAACAGCAGTCAAGTTGGTGGTTATTAACATATGTCTCTGGTACCAAGTCCCC	215
QY	181	AGGCAAAAGCCCCAAACTCATCTTTATAGGTCAGTATCGGCCCTCAGGGCTTTCTAA	240
Db	216	AGGCAAGGCCCCAAAACATCATGTTATAGGTCAGTATCGGCCCTCAGGGCTTTCTAA	275
QY	241	TTCGTTCTCTGAGGCTCCAAAGTCCTGGCAACAAGGCTCCCTGACATCTTGGGCTTCAGGC	300
Db	276	TTCGTTCTCTGAGGCTCCAAAGTCCTGGCAACAAGGCTCCCTGACATCTTGGGCTTCAGGC	335
QY	301	TGAGGACGAGGCTGATTATTACTGTGCTCATATACAAGATGTTCTCATGTCTTGG	360
Db	336	TGAGGACGAGGCTGATTATTACTGTGCTCATATACAAGATGTTCTCATGTCTTGG	395
QY	361	AACCTGGGACCAAGTCAACCTCCTAGGTACGCCCAAGCCAACTCCACTGTCACTGTGTT	420
Db	396	CGAGGGGACCAACTGACCTGAGTGTGAGCCAAAGGCTGCCCTTCGGTCACTGTGTT	455
QY	421	CCGCGCTCTCTGAGGAGTCCAAAGCCAAACAAGGCCACATAGTGTCTGATCACTGA	480
Db	456	CCGCGCTCTCTGAGGAGTCTCAAGCCAAACAAGGCCACATGTGTGTCTCATAGTGA	515
QY	481	CTTCTACCCGGGAGCTGTGACAGTGGCTCTGGAAGGCAATGCGACGCCCTGCAAGGCGGG	540
Db	516	CTTCTACCCGGGAGCGGTGACAGTGGCTCTGGAAGGCAATGAGCGGCCCTGCAAGGCGGG	575
QY	541	AGTGGAGACCAACCAACCTCCAAACAGAGCAACAACAAGTACGGGCGAGCACTACTCT	600
Db	576	AGTGGAGACCAACCAACCTCCAAACAGAGCAACAACAAGTACGGGCGAGCACTACTCT	635
QY	601	GAGCCTGACGCCCGAGCAAGTGAAGTCCCAAGAACTTACGCTTGCACGATCAGCATGA	660
Db	636	GAGCCTGACGCCCGAGCAAGTGAAGTCCCAAGAACTTACGCTTGCACGATCAGCATGA	695
QY	661	AGGAGACACCTGGATGTAAGACAGTGTCCC	691
Db	696	AGGAGACACCTGGGA-GAAGACAGTGTCCCC	725

RESULT 2  
US-08-378-939-11

; Sequence 11, Application US/08378939

PatentNo. 5876961  
GENERAL INFORMATION:  
APPLICANT: CROWE, JAMES SCOTT  
APPLICANT: LEWIS, ALAN PETER  
TITLE OF INVENTION: PRODUCTION OF ANTIBODIES  
NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS  
 ADDRESSEE: ROTHWELL  
 STREET: 555 THIRTEEN  
 CITY: WASHINGTON  
 STATE: D. C.  
 COUNTRY: U.S.  
 ZIP: 20004  
 COMPUTER READABLE FORM

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

URGENT APPLICATION DATA:

APPLICATION NUMBER: US/08/378,939  
FILING DATE:

CLASSIFICATION: 435

**RIOR APPLICATION**

APPLICATION NUMBER:  
 FILING DATE: 01-DEC-

FILED DATE: 01-DEC-1992  
ATTORNEY/AGENT INFORMATION:

NAME: ERNST, BARBARA G

REGISTRATION NUMBER: 30,  
REGISTRATION/DOCKET NUMBER:

REFERENCE/DOCUMENT NUMBER:  
ELECTRONIC COMMUNICATION INFORMATION

TELEPHONE: (202) 783-6040

TELEFAX: (202) 783-6031  
OPERATION FOR SEC ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 902 base pairs

TYPE: nucleic acid  
STANDARDNESS: both

STRANDEDNESS: DOCL  
TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO  
ATTI-SENSE: NO

INITIAL-SENSE: NO  
FEATURE:

NAME/KEY: CDS

LOCATION: 32..739

NAME/KEY: mat peptide

LOCATION: 89..739

NAME/REV: 010 0004170

NAME/RSI: 819\_pdrldue  
LOCATION: 32..86

378-939-11

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841

Local Similarity	92.5
Local Match	40.3

639; Conservative

1. **Introduction**

[illegible]

34 GGCCCTGGGCTCTGCTG

21 ECCCCCEAACEE

[illegible]

Query Match	40.9%;	Score 597.4;	DB 2;	Length 902;
Best Local Similarity	92.5%;	Pred. No. 2.2e-146;		
Matches 639; Conservative	0;	Mismatches 51;	Indels 1;	Gaps 1;

QY	1	GGCTGGGGCTCTGATTTCTCTCAACCTCTCACTCAAGGAGCAAGGGCTCTGAGGCCAGTC	60
Db	34	GGCTTGGGCTCTGGCTCTCTCAACCTCTCACTTGAACAAGGGCTCTGAGGCCAGTC	93
QY	61	TGCGCTGACTCAGTCTGCTCCGCTCGGTCTGGGTCTCTGAGACGTGATCAACATCTCTG	120
Db	94	TGCGCTGACTCAGCGCTCGCTCGGTCTGGGTCTCTGAGACGTGATCAACATCTCTG	153
QY	121	CATCGAACCCAGCAGTCAGTTGGTGTATTACTATGTCTCTGTATCCAAACAGCACCC	180
Db	154	CATCGAACCCAAAGATGTGGAGTTATACCTTGTCTCTGTATCCAGACAGCACCC	213
QY	181	AGGCAAGCCCCCAAACTCATTTATGAGGTCAGTATCGGCGCTCAGGGATTTCTAA	240
Db	214	AGGCAAGCCCCCAAAATCATTTATGAGGTCAGTAAAGGGCGCTCAAGGATTTCTAA	273
QY	241	TGCGTTCTTGCTCCAACTCTGGCAACAAGCGCTCCCTGACCATCTCTGGCTCCAGGC	300
Db	274	TGCGTTCTTGCTCCAACTCTGGCAACAAGCGCTCCCTGACCAATCTCTGGCTCCAGGC	333
QY	301	TGAGGACGAGGCGTGAATTTACTGTGTGCTCATATCAAGAAGTACTTCTCATGTCTTCG	360
Db	334	TGAGGACGAGGCGTGAATTTACTGTGTGCTCATATGCAAGTGTATCATGTGTCTTCG	393
QY	361	AATCGGACCAAGTCAACGCTCTTAGTCAAGGCCAAGGCCAACCCCATGTCTCTGTT	420
Db	394	CGAGGAGCAACAACGTGACGCTCTTAGTCAAGGCCAAGGCTGCCCTCTCGGTCTCTGTT	453
QY	421	CCGCGCTCTCTTGAGAGGCTCCAAAGCCAAAGGCCACACTAGTGTCTGATCAGTGA	480
Db	454	CCGCGCTCTCTTGAGAGGCTTCAAGCCAAAGGCCACACTGTGTGTCTCATATATGA	513



CURRENT FILING DATE: 2000-10-30  
NUMBER OF SEQ ID NOS: 1833  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 970  
LENGTH: 543  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-702-705-970

Query Match 32.9%; Score 480.6; DB 4; Length 543;  
Best Local Similarity 92.8%; Pred. No. 5.3e-116;  
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 100 ACAGTCGATCACCATCTCTGCTGAGACGAGCAGTCACTGATGTTATTAATATGT 159  
DB 543 ACAGTCGATCACCATCTCTGCTGAGACGAGCAGTCACTGATGTTATTAATATGT 484  
QY 160 CTCCTGTGACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 219  
DB 483 CTCCTGTGACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 424  
QY 220 TCGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 279  
DB 423 TCGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 364  
QY 280 GACCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 339  
DB 363 GACCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 304  
QY 340 AAGTACTTCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399  
DB 303 CAGTACTTCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 244  
QY 400 CAACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459  
DB 243 TCGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 184  
QY 460 ACTAGTGTCTGATCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519  
DB 183 ACTAGTGTCTGATCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 124  
QY 520 TCGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
DB 123 TAGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 64  
QY 580 GTACCGGCGCAGCAGCTACTGAGCTGAGCCTGAGCCTGAGCAGTGAAGTCCA 639  
DB 63 GTACCGGCGCAGCAGCTACTGAGCTGAGCCTGAGCCTGAGCAGTGAAGTCCA 4  
QY 640 CAG 642  
DB 3 CAG 1

RESULT 5  
US-09-736-457-970/C  
Sequence 970, Application US/09736457  
Patent No. 6509448  
GENERAL INFORMATION:  
APPLICANT: Wang, Tonglong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Veddyck, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fann, Lijun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.478CIS  
CURRENT APPLICATION NUMBER: US/09/736,457

CURRENT FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 1864  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 970  
LENGTH: 543  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-736-457-970

Query Match 32.9%; Score 480.6; DB 4; Length 543;  
Best Local Similarity 92.8%; Pred. No. 5.3e-116;  
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 100 ACAGTCGATCACCATCTCTGCTGAGACGAGCAGTCACTGATGTTATTAATATGT 159  
DB 543 ACAGTCGATCACCATCTCTGCTGAGACGAGCAGTCACTGATGTTATTAATATGT 484  
QY 160 CTCCTGTGACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 219  
DB 483 CTCCTGTGACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 424  
QY 220 TCGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 279  
DB 423 TCGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 364  
QY 280 GACCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 339  
DB 363 GACCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 304  
QY 340 AAGTACTTCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399  
DB 303 CAGTACTTCTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 244  
QY 400 CAACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459  
DB 243 TCGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 184  
QY 460 ACTAGTGTCTGATCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519  
DB 183 ACTAGTGTCTGATCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 124  
QY 520 TCGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
DB 123 TAGGACCTCAGGAGTTTCTTAATCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 64  
QY 580 GTACCGGCGCAGCAGCTACTGAGCTGAGCCTGAGCCTGAGCAGTGAAGTCCA 639  
DB 63 GTACCGGCGCAGCAGCTACTGAGCTGAGCCTGAGCCTGAGCAGTGAAGTCCA 4  
QY 640 CAG 642  
DB 3 CAG 1

RESULT 6  
US-09-614-124B-970/C  
Sequence 970, Application US/09614124B  
Patent No. 6630574  
GENERAL INFORMATION:  
APPLICANT: Wang, Tonglong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Veddyck, Tom  
APPLICANT: Carter, Darrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fann, Lijun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.478C9  
CURRENT APPLICATION NUMBER: US/09/614,124B  
CURRENT FILING DATE: 2001-07-11  
NUMBER OF SEQ ID NOS: 1668

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 970  
LENGTH: 543  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-614-124B-970

Query Match 32.9%; Score 480.6; DB 4; Length 543;  
Best Local Similarity 92.8%; Pred. No. 5.3e-116;  
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
QY 100 ACAGTCATGATCACCCTCTCTGCACTGAGACAGAGTCACTGGTGGTTATATCAATAT 159
Db 543 ACAGTCATGATCACCCTCTCTGCACTGAGACAGAGTCACTGGTGGTTATATCAATAT 484
QY 160 CTCCTGGTACCAACAGACAGACCCAGCAAGCCCCCAACTCATCTATTATAGAGTCA 219
Db 483 CTCCTGGTACCAACAGACAGACCCAGCAAGCCCCCAACTCATCTATTATAGAGTCA 424
QY 220 TCGGCCCCAGAGGGTTTCTAATGCTCTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 279
Db 423 TCGGCCCCAGAGGGTTTCTAATGCTCTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 364
QY 280 GACCATCTGAGGCTCCAGGCTGAGACAGAGCTGATTTATTAAGTCTGCTCATATACA 339
Db 363 GACCATCTGAGGCTCCAGGCTGAGACAGAGCTGATTTATTAAGTCTGCTCATATACA 304
QY 340 AAGTACTTCTCATGTTCTTGGAACTGGAGCAAGTCAAGCTCTAGTCAAGCCCAAGGC 399
Db 303 AAGTACTTCTCATGTTCTTGGAACTGGAGCAAGTCAAGCTCTAGTCAAGCCCAAGGC 244
QY 400 CAACCCCACTGCTCACTGTTCCCGCCCTCTGAGAGAGTCCAGGCCCAAGGCCAC 459
Db 243 TCGCCCCCTGCTCACTGTTCCCGCCCTCTGAGAGAGTCCAGGCCCAAGGCCAC 184
QY 460 ACTAGTGTGTCTATGATCACTGATCTTACCCGAGAGTGTGACAGTGGCTGGAAGCAGA 519
Db 183 ACTAGTGTGTCTATGATCACTGATCTTACCCGAGAGTGTGACAGTGGCTGGAAGCAGA 124
QY 520 TGGCAGCCCCGTCAAGCGGGAGTGAAGACCAACCAACCTTCCAAACAGAGCAACAA 579
Db 123 TGGCAGCCCCGTCAAGCGGGAGTGAAGACCAACCAACCTTCCAAACAGAGCAACAA 64
QY 580 GTACGCGGCGAGAGCTACTGAGCTGAGCGCCGAGAGAGTGGAGAGTCCCAAGAGCTA 639
Db 63 GTACGCGGCGAGAGCTACTGAGCTGAGCGCTGAGAGAGTGGAGAGTCCCAAGAGCTA 4
QY 640 CAG 642
Db 3 CAG 1
```

## RESULT 7

US-09-671-325-970/C  
Sequence 970, Application US/09671325  
Patent No. 6667154  
GENERAL INFORMATION:  
APPLICANT: Wang, Tonglong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Derrick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Pan, Liqun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.478C12  
CURRENT APPLICATION NUMBER: US/09/671,325  
CURRENT FILING DATE: 2000-09-26  
NUMBER OF SEQ ID NOS: 1825  
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 970  
LENGTH: 543  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-671-325-970

Query Match 32.9%; Score 480.6; DB 4; Length 543;  
Best Local Similarity 92.8%; Pred. No. 5.3e-116;  
Matches 504; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
QY 100 ACAGTCATGATCACCCTCTCTGCACTGAGACAGAGTCACTGGTGGTTATATCAATAT 159
Db 543 ACAGTCATGATCACCCTCTCTGCACTGAGACAGAGTCACTGGTGGTTATATCAATAT 484
QY 160 CTCCTGGTACCAACAGACAGACCCAGCAAGCCCCCAACTCATCTATTATAGAGTCA 219
Db 483 CTCCTGGTACCAACAGACAGACCCAGCAAGCCCCCAACTCATCTATTATAGAGTCA 424
QY 220 TCGGCCCCAGAGGGTTTCTAATGCTCTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 279
Db 423 TCGGCCCCAGAGGGTTTCTAATGCTCTCTGCTCTCAAGTCTGGCAACAGGCTCCCT 364
QY 280 GACCATCTGAGGCTCCAGGCTGAGACAGAGCTGATTTATTAAGTCTGCTCATATACA 339
Db 363 GACCATCTGAGGCTCCAGGCTGAGACAGAGCTGATTTATTAAGTCTGCTCATATACA 304
QY 340 AAGTACTTCTCATGTTCTTGGAACTGGAGCAAGTCAAGCTCTAGTCAAGCCCAAGGC 399
Db 303 AAGTACTTCTCATGTTCTTGGAACTGGAGCAAGTCAAGCTCTAGTCAAGCCCAAGGC 244
QY 400 CAACCCCACTGCTCACTGTTCCCGCCCTCTGAGAGAGTCCAGGCCCAAGGCCAC 459
Db 243 TCGCCCCCTGCTCACTGTTCCCGCCCTCTGAGAGAGTCCAGGCCCAAGGCCAC 184
QY 460 ACTAGTGTGTCTATGATCACTGATCTTACCCGAGAGTGTGACAGTGGCTGGAAGCAGA 519
Db 183 ACTAGTGTGTCTATGATCACTGATCTTACCCGAGAGTGTGACAGTGGCTGGAAGCAGA 124
QY 520 TGGCAGCCCCGTCAAGCGGGAGTGAAGACCAACCAACCTTCCAAACAGAGCAACAA 579
Db 123 TGGCAGCCCCGTCAAGCGGGAGTGAAGACCAACCAACCTTCCAAACAGAGCAACAA 64
QY 580 GTACGCGGCGAGAGCTACTGAGCTGAGCGCCGAGAGAGTGGAGAGTCCCAAGAGCTA 639
Db 63 GTACGCGGCGAGAGCTACTGAGCTGAGCGCTGAGAGAGTGGAGAGTCCCAAGAGCTA 4
QY 640 CAG 642
Db 3 CAG 1
```

## RESULT 8

US-08-487-550-9  
Sequence 9, Application US/08487550  
Patent No. 6113898  
GENERAL INFORMATION:  
APPLICANT: Anderson, Darrell R.  
TITLE OF INVENTION: MONOCLONAL ANTIBODIES SPECIFIC  
TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATED FORMS THEREOF,  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS  
TITLE OF INVENTION: IMMUNOSUPPRESSANTS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS





QY	121	CACGTGAACACAGAGTCAAGTGTGGGTATATACATATGTCTCCGTGACCAAGCACCC	180
Db	126	CATGTGGAGACCTCCAACTTGGAGTTAT---GATCTCAATTGGTACCAAGCTCCC	182
QY	181	AGGCAAAAGCCCCCAAATCATCATATTATAGGTGCGTAAATCGGCCCTCAGGGATTCTAA	240
Db	183	AGGAACGGCCCCCAAATCCTCATCATATGACATTAAACAAGGACCCCTCAGGAATTTCTGA	242
QY	241	TGCGCTTCTCTGGGTCCAAAGTCTGGCAACACGGCTCCCTGACACATCTCTGGGCTCCAGGC	300
Db	243	CCGATTTCTTGGTCTCCAAAGCTGGTACCGCGGCTCCTCCGACATCACTGGGCTCCAGAC	302
QY	301	TGAGAGACGAGCTGATTTATTACTGTGCTCATATACAAAGAA--TACTTCTCATGTCTT	357
Db	303	TGAGAGATGAGGCTGATTATTATCTGCCAGTCTTATACACAGAGCTGAAATGTCAGGATTT	362
QY	358	CGAAACTGGGACCAAGTCAACCGTCTAGGTGACGCCAAGGCCAACCCCACTGTCACTCT	417
Db	363	CGAGAGAGGGACCCGGCTGACCGTCTAGGTGACGCCAAGGCTGGCCCCCTCGGTCACTCT	422
QY	418	GTTTCCGGCCTCTCTCTGAGAGAGTCCAAACCAAGGCCACACTAGTGTGTGATCAG	477
Db	423	GTTTCCGCCCTCTCTCTGAGAGAGTTCCAAACCAAGGCCACACTGTGTGTCTCATPAG	482
QY	478	TGACTTTTACCCCGAGAGCTGTGACAGTGCCTGGAAGGACAGTGGACGCCCTGCATGGC	537
Db	483	TGACTTTTACCCCGAGAGCCGTGACAGTGCCTGGAAGGACAGTGCACGCCCTGCATAGGC	542
QY	538	GGGAGTGTGAGACCAACCAAAACCTTCCAAACAGACGACAAACAAGTACCGCGCCAGACACTA	597
Db	543	GGGAGTGTGAGACCAACCAACCTTCCAAACAGACGACAAACAAGTACCGCGGCCACAGACTA	602
QY	598	CTTGAAGCTTACGCCCCGAGCAGTGGAAATCTCCACGAAGCTTACAGCTTCCAGGTCAAGCA	657
Db	603	CTTGAAGCTTACGCCCTTGAAGCAGTGGAAATCTCCACGAAGCTTACAGCTTCCAGGTCAAGCA	662
QY	658	TGAAGGGAGCAGCGTGAATGTAAGACAGTGTGCC	691
Db	663	TGAAGGGAGCAGCGTGA--GAAACAGTGTGCC	695

RESULT 10  
US-03-049-672A-25  
Sequence 25, Application US/09049672A  
Patent No. 6135941  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Tang, Y. Tom  
APPLICANT: Yue, Henry  
APPLICANT: Au-Young, Janice  
APPLICANT: Corley, Neil C  
APPLICANT: Guegler, Karl J.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/049,672A  
FILING DATE: HERewith  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:

```

1  APPLICATION NUMBER:
2  FILING DATE:
3  ATTORNEY/AGENT INFORMATION:
4  NAME: Cerrone, Michael C
5  REGISTRATION NUMBER: 39,132
6  REFERENCE/DOCKET NUMBER: PF-0497 US
7  TELECOMMUNICATION INFORMATION:
8  TELEPHONE: 650-855-0555
9  TELEFAX: 650-845-4166
10
11  TELEX:
12
13  INFORMATION FOR SEQ ID NO: 25:
14  SEQUENCE CHARACTERISTICS:
15  LENGTH: 895 base pairs
16  TYPE: nucleic acid
17  STRANDEDNESS: single
18  TOPOLOGY: linear
19  IMMEDIATE SOURCE:
20  LIBRARY: LUNGUT13
21  CLONE: 3116314
22  US-09-049-672A-25

```

[illegible]



```

Db      75 AGGGCTGACTCAGACCCCTCGGTGTCAGAGAGCTTGAGACAGACCGCACACTCAGCTG 134
QY      121 CACTGGAACCCAGCAGCTGCTGGTGTATTAATGCTCTCTCTGTCACCAAGCACCC 180
Db      135 CACCGGAAACA--ACAAATGTGGCACCAAGAGAGCTTGCTGACAGACACCA 191
QY      181 AGGCAAGAGCCCAAACTCATATTATGAGGTCAAGTAATGGCCCTCAGGGGTTTCTAA 240
Db      192 GGGGCAACCTCCCAAACTCTCTCTACAGAAATATAACGGCCCTCAGGATCTCAGA 251
QY      241 TCGCTTCTGTGCTCCAACTGCGCAACAGGCTCTCCCTGACCAATCTGTGGCTTCAGGC 300
Db      252 GAGATTATCTGATCCAGGTCTCAGAGCCACATCTCTCCGACATTAATCTGATCTCAAGCC 311
QY      301 TGAGAGCAGAGGTGATTATTAATGCTGCTCATATTAAGAAAG---TACTTTCATGCTT 357
Db      312 TGAGAGCAGAGGTGATTATTAATGCTGCTCATATTAAGAAAG---TACTTTCATGCTT 371
QY      358 CGGAACTGGGACCAAGGTACCGCTCTAGCTAGGCCCAAGGCCAACCCCACTGCTACTCT 417
Db      372 CGGCGAGGGACCAAGCTACCGCTCTAGCTAGGCCCAAGGCCAACCCCACTGCTACTCT 431
QY      418 GTTCCGCCCTCTCTCTGAGAGAGTCCAAAGCCCAAGGCCCAACTGATGCTGATCAG 477
Db      432 GTTCCCACTCTCTCTGAGAGAGTCCAAAGCCCAAGGCCCAACTGATGCTGATCAG 491
QY      478 TGACTTCTACCCGAGAGCTGTGACAGTGGCTCTGAGAGGACATGAGCAGCCCGTCAAGGC 537
Db      492 TGACTTCTACCCGAGAGCTGTGACAGTGGCTCTGAGAGGACATGAGCAGCCCGTCAAGGC 551
QY      538 GGGAGTGGAGCAACCAAACTCTCCAAACAGACCAACAAATGACGGCGCCAGACGCTA 597
Db      552 GGGAGTGGAGCAACCAACCTCTCCAAACAGACCAACAAATGACGGCGCCAGACGCTA 611
QY      598 CCGAGCTCTGAGCCCGGACAGTGGAGTCCCAAGAGCTACAGCTGCAAGTCAAGCA 657
Db      612 CCGAGCTCTGAGCCCGGACAGTGGAGTCCCAAGAGCTACAGCTGCAAGTCAAGCA 671
QY      658 TGAAGGAGCACCCTGATGTAAGACAGTGTGCC 691
Db      672 TGAAGGAGCACCCTGATGTAAGACAGTGTGCC 704

RESULT 13
US-09-152-060-47
; Sequence 47, Application US/09152060
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FID REFERENCE: P2003P1.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/046,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19

```

```

; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 885
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-152-060-47

Query Match      28.6%; Score 418; DB 4; Length 885;
Best Local Similarity 78.1%; Pred. No. 1,3e-99;
Matches 542; Conservative 0; Mismatches 145; Indels 7; Gaps 3;

QY      1 GGGCTGGGCTCTGCTATTCCTCAACCTCTCTCACTGAGGACAGAGGCTCGGGCCAGTC 60
Db      23 GCGCTGGGCTCTGCTCTCTCTCTGACCTCTCTCACTGAGGAGTCAAGTGTCCAGG 82
QY      61 TGCCCTGACTCACTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db      83 AGGGCTGACTCAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 142
QY      121 CACTGGAACCCAGCAGCTGCTGGTGTATTAATGCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db      143 CACCGGAAACA--ACAAATGTGGCACCAAGAGAGCTTGCTGACAGACCA 199
QY      181 AGGCAAGAGCCCAAACTCATATTAAGAGTCAAGTAATGGCCCTCAGGGTTTCTAA 240
Db      200 GGGCCACCTCTCCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 259
QY      241 TCGCTTCTGTGCTCTCAAGTCTGGCAACGGCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db      260 GAGATTATCTGATCCAGGTCAGAGCCACATCTCTCTCTCTCTCTCTCTCTCTCTCT 319
QY      301 TGAGAGCAGAGGTGATTATTAATGCTGCTCATATTAAGAAAG---TACTTTCATGCTT 357
Db      320 TGAGAGCAGAGGTGATTATTAATGCTGCTCATATTAAGAAAG---TACTTTCATGCTT 379
QY      358 CGGAACTGGGACCAAGTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 417
Db      380 CGGAGTGGAGCAACCACTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 439
QY      418 GTTCCGCCCTCTCTCTGAGAGTCTCAAGGCAACAGGCAACAGCTGATGCTGATCAG 477
Db      440 GTTCCCACTCTCTCTGAGAGTCTCAAGGCAACAGGCAACAGCTGATGCTGATCAG 499
QY      478 TGACTTCTACCCGAGAGCTGTGACAGTGGCTCTGAGAGGACATGAGCAGCCCGTCAAGGC 537
Db      500 TGACTTCTACCCGAGAGCTGTGACAGTGGCTCTGAGAGGACATGAGCAGCCCGTCAAGGC 559
QY      538 GGGAGTGGAGCAACCAAACTCTCCAAACAGACCAACAAATGACGGCGCCAGACGCTA 597
Db      560 GGGAGTGGAGCAACCAACCTCTCCAAACAGACCAACAAATGACGGCGCCAGACGCTA 619
QY      598 CCGAGCTCTGAGCCCGGACAGTGGAGTCCCAAGAGCTACAGCTGACGCTCAGGTCAGCA 657
Db      620 CCGAGCTCTGAGCCCGGACAGTGGAGTCCCAAGAGCTACAGCTGACGCTCAGGTCAGCA 679
QY      658 TGAAGGAGCACCCTGATGTAAGACAGTGTGCC 691
Db      680 TGAAGGAGCACCCTGATGTAAGACAGTGTGCC 712

RESULT 14
US-08-523-894-5
; Sequence 5, Application US/08523894
; Patent No. 6136310
; GENERAL INFORMATION:
; APPLICANT: Hanna, Nabil
; APPLICANT: Newman, Roland A.
; APPLICANT: Reff, Mitchell E.
; TITLE OF INVENTION: Recombinant Anti-CD4 Antibodies for Human
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314-3187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/523,894  
FILING DATE: 06-SEP-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-165  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 702 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: lambda variable and constant domains in  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..702  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..702  
US-08-523-894-5

Query Match 28.4%; Score 414; DB 3; Length 702;  
Best Local Similarity 78.0%; Pred. No. 1.3e-98;  
Matches 341; Conservative 0; Mismatches 140; Indels 13; Gaps 3;

QY 1 GGCCCTGGGCTCTGCTATTCCTCAACCTCTCACTAGGACAGAGGCTCTGGGCCAGTC 60  
DB 3 GGCCCTGGGCTCTGCTCTGCTCTCGGCTCTCTGCTCACTTACAGACTCTGGGCTCTCA 62  
QY 61 TGGCCCTGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 63 TGAGTTGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122  
QY 121 CACTGGAACCAAGACAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
DB 123 TGGGGGAGAAC-----GTGGAAGAAAGTAGTAGTGTACACAGAGGC 173  
QY 181 AGGCAAAAGCCCCCAAACTATCTTTATGAGTCACTATCGGCTCTCAGGGGTTTCTTA 240  
DB 174 ACCGCAAGCCCCCTGTGTGCTATCTATGCTGACAGGAGCCCTCGGAGTCCCTGCG 233  
QY 241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 234 GCGATTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 293  
QY 301 TGAAGACGAGGCTGATTATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357  
DB 294 CCGGGATGAGGCTGACTATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 353  
QY 358 CGGAAGTGGAGCAAGGTCACGCTAGGTCAGCCCAAGGCAAGGCAAGGCAAGGCAAGGCA 417  
DB 354 CCGGAGAGGAGCAAGGTCAGCCTAGGTCAGCCCAAGGCAAGGCAAGGCAAGGCAAGGCA 413

QY 418 GTTCCCGCCCTCTCTGAGAGAGCTCCAGACCAAGGCCACAGTATGCTGATCAG 477  
DB 414 GTTCCCGCCCTCTCTGAGAGAGCTTCAAGCAACAGGCCACAGTATGCTGATCAG 473  
QY 478 TGACTTACCCGAGAGCTGTGACAGTGGCTGGAAGGACAGATGAGACCCGCTCAAGGC 537  
DB 474 TGACTTACCCGAGAGCCGTGACAGTGGCTGGAAGGACAGATGAGACCCGCTCAAGGC 533  
QY 538 GGGAGTGAAGACCAACCAACCTTCCAAACAGACCAACCAATATAGCGGCGACAGCTA 597  
DB 534 GGGAGTGAAGACCAACCAACCTTCCAAACAGACCAACCAATATAGCGGCGACAGCTA 593  
QY 598 CTTGAGCCTGACGCCCCGAGCAGTGAAGTCCACAGAGCTACAGCTGCGAGTCAAGCA 657  
DB 594 CTTGAGCCTGACGCGCTGAGCAGTGAAGTCCACAGAGCTACAGCTGCGAGTCAAGCA 653  
QY 658 TGAAGGAGCAAGCTGTGATGAAGACAGTGTCC 691  
DB 654 TGAAGGAGCAAGCTGTGAGTGAAGAGAGTGGCCCC 686

RESULT 15  
US-08-487-550-1  
Sequence 1, Application US/08487550  
Patent No. 611898  
GENERAL INFORMATION:  
APPLICANT: Anderson, Darrell R.  
TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC  
TO HUMAN B7.1 AND/OR B7.2 PRIMATED FORMS THEREOF,  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS  
NUMBER OF INVENTION: IMMUNOSUPPRESSANTS"  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: 699 Prince Street  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,550  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 012712-131  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6620  
TELEFAX: 703-836-2021  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 705 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..705  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 1..705  
US-08-487-550-1

Query Match 28.1%; Score 410.8; DB 3; Length 705;  
Best Local Similarity 77.7%; Pred. No. 9e-98;

Matches 539; Conservative 0; Mismatches 142; Indels 13; Gaps 3;

```
QY      1 GGCCTGGGCTCTGCTATTCTCTCACTCTCACTCAAGGACAGGATCTGGGCCAGTC 60
Db      6 GGTCCCCGCTCAGCTCTGGGCTCTGCTCTGCTCTGCTCCAGGTGACAGATGTGCTTA 65
QY     61 TGGCTGACTCAGTTCGCTCCGTCGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 120
Db     66 TGAATGACTCAGTTCGCTCCGTCGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 125
QY     121 CACTGGAACAGCAGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db     126 TGGGGAGAGACAGTA-----GAATGATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 176
QY     181 AGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db     177 AGGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 236
QY     241 TCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db     237 GCGATTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 296
QY     301 TGAAGACAGAGCTGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
Db     297 CGGGGATGAGGCTGACTATTACTGTGAGGTGTGAGGTGTGAGGTGTGAGGTGTGAGGTGT 356
QY     358 CGGAATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 417
Db     357 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416
QY     418 GTTCCCGCCCTCTCTGAGGAGCTCCAGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
Db     417 GTTCCCGCCCTCTCTGAGGAGCTCCAGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476
QY     478 TGAATCTTACCCGGGAGCTGTGACAGTGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 537
Db     477 TGAATCTTACCCGGGAGCTGTGACAGTGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 536
QY     538 GGAAGTGAAGACCAACCAACCTCCAAACAGAGACCAACCAACCAACCAACCAACCAACCA 597
Db     537 GGAAGTGAAGACCAACCAACCTCCAAACAGAGACCAACCAACCAACCAACCAACCAACCA 596
QY     598 CTTGAGCTGAGCGCCGAGAGAGTGAAGTCCACAGAGAGTGAAGTCCACAGAGTGAAGT 657
Db     597 CTTGAGCTGAGCGCCGAGAGAGTGAAGTCCACAGAGAGTGAAGTCCACAGAGTGAAGT 656
QY     658 TSAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 691
Db     657 TSAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 689
```

Search completed: April 26, 2004, 20:28:40  
Job time: 124.019 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 26, 2004, 15:46:02 ; Search time 3550.63 Seconds  
(without alignments)  
12279.158 Million cell updates/sec

Title: US-10-076-747-8  
Perfect score: 1460  
Sequence: 1 ggcctcggcgcctcgtatcc.....ggcagcagcagccacgcagct 1460

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	638.8	43.8	889	12	BG756342
2	633.6	43.3	786	12	BG758901
3	631.8	43.3	911	13	BQ712542
4	617.6	42.3	829	12	BG685732

5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	
615.6	614.8	612.4	610.8	607.8	606.8	606.4	606.4	604.4	603.8	602.8	602.8	601.6	601.6	599.2	598.8	598.2	598.2	597.6	597.6	596.4	595.4	595.2	595.2	594.8	593.2	593.2	593.2	591.4	591.4	590.6	590.2	589.8	589.2	589.2	589.2	588.4	587.8	587.4	587.2	586.4	
42.2	42.1	42.0	41.9	41.6	41.6	41.5	41.5	41.4	41.4	41.3	41.2	41.2	41.0	41.0	41.0	41.0	41.0	40.9	40.9	40.8	40.8	40.8	40.8	40.7	40.6	40.6	40.6	40.5	40.5	40.4	40.4	40.4	40.4	40.3	40.3	40.2	40.2	40.2	40.2	40.2	
868	810	729	729	1049	686	1150	1201	1201	840	903	945	1040	1063	739	813	953	958	1019	767	884	892	892	892	898	898	898	895	854	854	876	772	786	786	786	786	880	819	931	856	10	
BG483745	BG754419	BG754419	EM007694	EM920304	BG759257	BX446259	BX449367	BX462166	BG397282	BE561081	BX446252	BX410975	BM920655	BE397699	BE974961	BE976229	BX426264	BE271225	AV697043	BM051870	B1227090	B1820758	BQ711473	BE562370	BM918781	BE559866	BE975248	B1227345	BM051654	BE397742	B1224785	B1823305	B1823305	B1226887	BG756874	CD691527	BE975970	BE560212	BM051920	BE561313	
BG483745	602503383	602710036	603617126	AGENCOURT	602710936	BX446259	BX449367	BX462166	BG397282	602439065	BX446252	BX410975	BM920655	BE397699	602244532	602245104	BX426264	600943893	AV697043	603638744	602952652	603034354	AGENCOURT	601344455	AGENCOURT	601345771	602244754	602244754	603638839	601289887	602949465	603042121	602952002	602710363	EST8050	602246174	601345193	603638812	601344452		

## ALIGNMENTS

RESULT 1  
BG756342 889 bp mRNA EST 15-MAY-2001  
LOCUS 602713662P1 NIH\_MGC\_48 Homo sapiens CDNA clone IMAGE:4853919 5',  
DEFINITION mRNA sequence.  
ACCESSION BG756342  
VERSION BG756342.1 GI:14066995  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 889)  
NIH-MGC <http://imgc.ncbi.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rudin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.llnl.gov>  
Plate: LNCM1700 row: 1 column: 16  
High quality sequence stop: 810.

FEATURES  
source

Location/Qualifiers  
1. 889  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4853450"  
/issue\_type="primary B-cells from tonsils (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 48"  
/note="Organ: B-cells; Vector: pOT87; Site: 1. XhoI; Site: 2. EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 43.8%; Score 638.8; DB 12; Length 889;  
Best Local Similarity 95.4%; Pred. No. 2.1e-135;  
Matches 658; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

QY 1 GGCTGGGCTCTGCTATCTCTCACTCTCTCACTCAAGGACAGGAGTCTGGGCTCAATC 60
DB 34 GGCTGGGCTCTGCTCTCTCTCTCACTCTCTCACTCAAGGACAGGAGTCTGGGCTCAATC 93
QY 61 TGCCCTGACAGCTGAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 120
DB 94 TGCCCTGACAGCTGAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 153
QY 121 CACTGAGACAGAGTCAAGTCTGAGTCTGATTAATGATGATGATGATGATGATGATGATGAT 180
DB 154 CACTGAGACAGAGTCAAGTCTGAGTCTGATTAATGATGATGATGATGATGATGATGATGAT 213
QY 181 AGGCAAGGACCCCAACTATCATTTATAGGTCAATATGAGGCTCTGAGGCTCTGAGGCTCT 240
DB 214 AGGCAAGGACCCCAACTATCATTTATAGGTCAATATGAGGCTCTGAGGCTCTGAGGCTCT 273
QY 241 TGCCCTGACAGCTGAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 300
DB 274 TGCCCTGACAGCTGAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 333
QY 301 TGAGACAGAGTCTGATTTACTGCTGCTCATATATACAGAGTCTTCTCATGCTTCTGG 360
DB 334 TGAGACAGAGTCTGATTTACTGCTGCTCATATATACAGAGTCTTCTCATGCTTCTGG 393
QY 361 AACTGAGACAGAGTCAAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 420
DB 394 AACTGAGACAGAGTCAAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 453
QY 421 CCGGCTCTCTGAGAGTCTCAAGCAAGGACAGGCTAGTGTCTGATCAGTGA 480
DB 454 CCGGCTCTCTGAGAGTCTCAAGCAAGGACAGGCTAGTGTCTGATCAGTGA 513
QY 481 CTCTCTCTGAGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 540
DB 514 CTCTCTCTGAGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 573
QY 541 AGTGAAGACCAACCTCTCAACAGAGCAACAAAGTACAGGCTAGTCTGATCAGTGA 600
DB 574 AGTGAAGACCAACCTCTCAACAGAGCAACAAAGTACAGGCTAGTCTGATCAGTGA 633
QY 601 GAGCTGAGAGGCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAG 660
DB 634 GAGCTGAGAGGCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAG 693
QY 661 AGGAGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690
DB 694 AGGAGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 723

```

## RESULT 2

BG758901

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1. 786

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4853450"

/issue\_type="primary B-cells from tonsils (cell line)"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 48"

/note="Organ: B-cells; Vector: pOT87; Site: 1. XhoI; Site: 2. EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected &gt;500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC library."

Query Match 43.3%; Score 632.6; DB 12; Length 786;

Best Local Similarity 95.7%; Pred. No. 5.2e-134;

Matches 661; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

1 GGCTGGGCTCTGCTATCTCTCACTCTCTCACTCAAGGACAGGAGTCTGGGCTCAATC 60

34 GGCTGGGCTCTGCTCTCTCTCTCACTCTCTCACTCAAGGACAGGAGTCTGGGCTCAATC 93

61 TGCCCTGACAGCTGAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 120

94 TGCCCTGACAGCTGAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 153

121 CACTGAGACAGAGTCAAGTCTGAGTCTGATTAATGATGATGATGATGATGATGATGATGAT 180

154 CACTGAGACAGAGTCAAGTCTGAGTCTGATTAATGATGATGATGATGATGATGATGATGAT 213

181 AGGCAAGGACCCCAACTATCATTTATAGGTCAATATGAGGCTCTGAGGCTCTGAGGCTCT 240

214 AGGCAAGGACCCCAACTATCATTTATAGGTCAATATGAGGCTCTGAGGCTCTGAGGCTCT 273

241 TGCCCTGACAGCTGAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 300

274 TGCCCTGACAGCTGAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 333

301 TGAGACAGAGTCTGATTTACTGCTGCTCATATATACAGAGTCTTCTCATGCTTCTGG 360





DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM1624 row: f column: 08  
 High quality sequence stop: 821.  
 Location/Qualifiers

## FEATURES

source

1..829  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4765447"  
 /tissue\_type="primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_48"  
 /note="Organ: B-cells; Vector: pOT81; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC library."

## ORIGIN

Query Match

Best Local Similarity 42.3%; Score 617.6; DB 12; Length 829;

Matches 65; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 1 GGCTTGGGCTCTGCTATTCCTCAACCTCCTCACTCAAGGACAGGGCTCTGGGCCCAATC 60  
 DB 34 GGCTTGGGCTCTGCTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 93  
 QY 61 TGGCTTGGGCTCTGCT 120  
 DB 94 TGGCTTGGGCTCTGCT 153  
 QY 121 CACTGGAACGAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 DB 154 CACTGGAACGAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 213  
 QY 181 AGGCAAAAGCCCAAACTCACTTATTAAGGTCATTAATGCTGCTGCTGCTGCTGCTGCTGCT 240  
 DB 214 AGGCAAAAGCCCAAACTCACTTATTAAGGTCATTAATGCTGCTGCTGCTGCTGCTGCTGCT 273  
 QY 241 TGGCTTGGGCTCTGCTCAAGTCTGCAACAGGCTCTCTGACCATCTCTGCTCTCAAGGCT 300  
 DB 274 TGGCTTGGGCTCTGCTCAAGTCTGCAACAGGCTCTCTGACCATCTCTGCTCTCAAGGCT 333  
 QY 301 TGAAGAGAGGCTGATTATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 DB 334 TGAAGAGAGGCTGATTATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 393  
 QY 361 AACTGGAGCAAGGTCACTCTAGGTCAAGCCCAAGGCAAGCCCAAGGCAAGCCCAAGGCAAG 420  
 DB 394 AACTGGAGCAAGGTCACTCTAGGTCAAGCCCAAGGCAAGCCCAAGGCAAGCCCAAGGCAAG 453  
 QY 421 CCGGCTCTCTCTGAGAGGCTCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 480  
 DB 454 CCGGCTCTCTCTGAGAGGCTCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 513  
 QY 481 CTTCTACCGGAGAGCTGTGACAGTGGGCTGGAAGGAGATGGAGGAGCCGCTCAAGGCGGG 540  
 DB 514 CTTCTACCGGAGAGCTGTGACAGTGGGCTGGAAGGAGATGGAGGAGCCGCTCAAGGCGGG 573  
 QY 541 AGT-GGAGACCAACCAACCTCCAAACAGAGCAACCAACCAAGTACGGGGCAGAGAGTACC 599  
 DB 574 AGTGGAGACCAACCAACCTCCAAACAGAGCAACCAACCAAGTACGGGGCAGAGAGTACC 633  
 QY 600 TGAAGCTGAGCGCCGAGAGTGAAGTCCAGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCT 659  
 DB 634 TGAAGCTGAGCGCCGAGAGTGAAGTCCAGAGGCTCAAGGCTCAAGGCTCAAGGCTCAAGGCT 693

QY 660 AAGGAGACACCGTGGATGAAGACAGTGTCTCC 691  
 DB 694 AAGGAGACACCGTGGATGAAGACAGTGTCTCC 725

## FEATURES

source

RESULT 5  
 LOCUS BG483745  
 DEFINITION 60250383F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4616829 5',  
 mRNA sequence.  
 ACCESSION BG483745  
 VERSION BG483745.1 GI:13415934  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 868)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM1372 row: e column: 22  
 High quality sequence stop: 834.  
 Location/Qualifiers

## ORIGIN

source

1..868  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4616829"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_id="NIH\_MGC\_77"  
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site\_1:  
 SfiI (ggccgctggcc); Site\_2: SfiI (ggccatgctggc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CAAGGCAATATGAC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCGGAGCGGCGGCAATG-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.9  
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH\_MGC library."

Query Match

Best Local Similarity 42.2%; Score 615.6; DB 12; Length 868;

Matches 66; Conservative 0; Mismatches 49; Indels 4; Gaps 2;

QY 1 GGCTTGGGCTCTGCTATTCCTCAACCTCCTCACTCAAGGCAAGGCTCTGGGCCCAATC 60  
 DB 42 GGCTTGGGCTCTGCTATTCCTCAACCTCCTCACTCAAGGCAAGGCTCTGGGCCCAATC 101  
 QY 61 TGGCTTGGGCTCTGCT 120  
 DB 102 TGGCTTGGGCTCTGCT 161  
 QY 121 CACTGGAACGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 DB 162 CACTGGAACGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 221  
 QY 181 AGGCAAAAGCCCAAACTCACTTATTAAGGTCATTAATGCTGCTGCTGCTGCTGCTGCTGCT 240  
 DB 222 AGGCAAAAGCCCAAACTCACTTATTAAGGTCATTAATGCTGCTGCTGCTGCTGCTGCTGCT 281

QY 241 TGGCTTCTGCTGCAAGTCTGGCAACAGGCGCTCCCTGACCATCTGCTGGGCTCCAGGC 300  
 DB 282 TGGCTTCTGCTGCTGCAAGTCTGGCAACAGGCGCTCCCTGACCATCTGCTGGGCTCCAGGC 341  
 QY 301 TGAAGACGAGGCTGATTATTACTGCTGCTCATATACAAAGATCTTCTCATATGCTTCGG 360  
 DB 342 TGAAGACGAGGCTGATTATTACTGCTGCTCATATACAAAGATCTTCTCATATGCTTCGG 401  
 QY 361 AACTGGACACCAAGTCAACCGTCTGATGTCAGCCCAAGCCCAACCCCACTGTCATCTGCTT 420  
 DB 402 AACTGGACACCAAGTCAACCGTCTGATGTCAGCCCAAGCCCAACCCCACTGTCATCTGCTT 461  
 QY 421 CCGCGCTCTCTGCTGAGAGCTCCAGCCCAAGCCCAACCTAGTGTCTGATGATGTA 480  
 DB 462 CCGCGCTCTCTGAGAGCTCCAGCCCAAGCCCAACCTAGTGTCTGATGATGTA 521  
 QY 481 CTCTTACCCGGAGCTGTGACAGTGGCTTGAAGGACAGTGGCAAGCCCGCTCAAGCGGG 540  
 DB 522 CTCTTACCCGGAGCTGTGACAGTGGCTTGAAGGACAGTGGCAAGCCCGCTCAAGCGGG 581  
 QY 541 AGTGAAGACCAACCAACCTCCCAACAGAGCAACCAAGTACGGGGCAGAGCTACT 600  
 DB 582 AGTGAAGACCAACCAACCTCCCAACAGAGCAACCAAGTACGGGGCAGAGCTACT 641  
 QY 601 GAGCCTGACGCGCCGAGCA--GTGAAATCCACAGAAAGCTACAGCTGCAGGTCAAGC-- 656  
 DB 642 GAGCCTGACGCGCCGAGCAAGTGAAGAAAGTCCACAGAGCTACAGCTGCAGGTCAAGC-- 701  
 QY 657 ATGGAAGGACACCGTGTGATGAAGACAGTGTGCTCCCTACAGAAAGTCTTATAGTTCTC 716  
 DB 702 TTGAAGGAGGACACCGTGTGATGAAGACAGTGTGCTCCCTACAGAAAGTCTTATAGTTCTC 761  
 QY 717 TA 718  
 DB 762 AA 763

RESULT 6  
 LOCUS BG754419 810 bp mRNA linear EST 15-MAY-2001  
 DEFINITION 602710036F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4846521 5',  
 mRNA sequence.  
 ACCESSION BG754419  
 VERSION BG754419.1 GI:14065072  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 810)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Louis W. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ULNL at:  
 http://image.jnl.gov  
 Plate: LCM1687 row: h column: 10  
 High quality sequence step: 805.  
 Location/Qualifiers  
 1..810  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4846521"  
 /issue\_type="Primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_48"

## FEATURES

source  
 1..810  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4846521"  
 /issue\_type="Primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH\_MGC\_48"

/note="Organ: B-cells; Vector: pOT7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC library."

Query Match 42.1%; Score 614.8; DB 12; Length 810;  
 Best Local Similarity 96.1%; Pred. No. 6.3e-130;  
 Matches 663; Conservative 0; Mismatches 22; Indels 5; Gaps 3;

QY 1 GGCCTGGGCTCTGCTATTCCTCAACCTCCCTCACTCAGGCGACAGGGTCCGGGCCAGATC 60  
 DB 18 GGCCTGGGCTCTGCTATTCCTCAACCTCCCTCACTCAGGCGACAGGGTCCGGGCCAGATC 76  
 QY 61 TGGCTGACTCAGTCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 DB 77 TGGCTGACTCAGTCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 136  
 QY 121 CACTGGAACCAAGATGACGTGTGTGTATTAATATATGCTCTGCTGCTGCTGCTGCTGCT 180  
 DB 137 CACTGGAACCAAGATGATGTTGGAGTTATTAACCTTGCTCTGCTGCTGCTGCTGCTGCT 196  
 QY 181 AGGCAAGCCCCCAACTCATTTATGAGTGCAGTAATCGGCGCTCAGGGGCTTTCTAA 240  
 DB 197 AGGCAAGCCCCCAACTCATTTATGAGTGCAGTAATCGGCGCTCAGGGGCTTTCTAA 256  
 QY 241 TGGCTTCTGCTGCTCAAGTCTGCAACAGGCGCTCTCTGACCATCTCTGAGGCTTCAAGC 300  
 DB 257 TGGCTTCTGCTGCTCAAGTCTGCAACAGGCGCTCTCTGACCATCTCTGAGGCTTCAAGC 316  
 QY 301 TGAAGACGAGGCTGATTATTACTGCTGCTCATATACAGAAAGTCTTATGATGCTTCGG 360  
 DB 317 TGAAGACGAGGCTGATTATTACTGCTGCTCATATACAGAAAGTCTTATGATGCTTCGG 373  
 QY 361 AACTGGACACCAAGTCAACCGTCTGATGTCAGCCCAAGCCCAACCCCACTGTCATCTGTT 420  
 DB 374 AACTGGACACCAAGTCAACCGTCTGATGTCAGCCCAAGCCCAACCCCACTGTCATCTGTT 433  
 QY 421 CCGCGCTCTCTGAGAGGCTCCAGCCCAAGCCCAACCTAGTGTCTGATGATGTA 480  
 DB 434 CCGCGCTCTCTGAGAGGCTCCAGCCCAAGCCCAACCTAGTGTCTGATGATGTA 493  
 QY 481 CTCTTACCCGGAGCTGTGACAGTGGCTTGAAGGACAGTGGCGCCGCTCAAGGCGGG 540  
 DB 494 CTCTTACCCGGAGCTGTGACAGTGGCTTGAAGGACAGTGGCGCCGCTCAAGGCGGG 553  
 QY 541 AGTGAAGACCAACCAACCTCCCAACAGAGCAACCAAGTACGGGGCCAGAGCTTACT 600  
 DB 554 AGTGAAGACCAACCAACCTCCCAACAGAGCAACCAAGTACGGGGCCAGAGCTTACT 613  
 QY 601 GAGCTTGAAGCCCGGAGAGGTAAGTCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 660  
 DB 614 GAGCTTGAAGCCCGGAGAGGTAAGTCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 673  
 QY 661 AGGAGACACCGTGTGATGAAGAGAGTGTCC 690  
 DB 674 AGGAGACACCGTGTGA-GAAGACAGTGGGCC 702

RESULT 7  
 LOCUS BF128674 729 bp mRNA linear EST 24-OCT-2000  
 DEFINITION 601810939F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4054065 5',  
 mRNA sequence.  
 ACCESSION BF128674  
 VERSION BF128674.1 GI:10967714  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 729)  
 NIH-MGC http://mgs.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LNCM893 row: m column: 10  
 High quality sequence stop: 696.  
 Location/Qualifiers  
 1..729  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4054065"  
 /tissue\_type="primary B-cells from tonsils (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH MGC 48"  
 /note="Organ: B-cells; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming;  
 directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCAAGAG(G). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by Ling  
 Hong in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 42.0%; Score 613; DB 10; Length 729;  
 Best Local Similarity 96.7%; Pred. No. 1.6e-129;  
 Matches 670; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

1 GGCCTGGGCTCTGTATTCCTCAACCTCCTCACTCAAGGACAGAGGCTCTGGGCCCAATC 60  
 1 GGCTGGGCTCTGTATTCCTCAACCTCCTCACTCAAGGACAGAGGCTCTGGGCCCAATC 60  
 61 TGCCCTGACTCAAGTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 61 TGCCCTGACTCAAGTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 121 CACTGGAACCAAGCAAGTCAAGTGTGTTAACTATGTCCTCTGTTACCAACAGCAACC 180  
 121 CACTGGAACCAAGCAAGTGTGTTAACTATGTCCTCTGTTACCAACAGCAACC 180  
 121 CACTGGAACCAAGCAAGTGTGTTAACTATGTCCTCTGTTACCAACAGCAACC 180  
 121 CACTGGAACCAAGCAAGTGTGTTAACTATGTCCTCTGTTACCAACAGCAACC 180  
 181 AGGCAAGAGCCCAACATCATTTATGAGGTCAAGTCAAGGCTCAAGGCTTCTCTA 240  
 181 AGGCAAGAGCCCAACATCATTTATGAGGTCAAGTCAAGGCTCAAGGCTTCTCTA 240  
 181 AGGCAAGAGCCCAACATCATTTATGAGGTCAAGTCAAGGCTCAAGGCTTCTCTA 240  
 241 TGCGTTCTGCTGCTCAAGTCTGGCAACAGGCTCCCTGACCATCTTGGGCTCCAGGC 300  
 241 TGCGTTCTGCTGCTCAAGTCTGGCAACAGGCTCCCTGACCATCTTGGGCTCCAGGC 300  
 241 TGCGTTCTGCTGCTCAAGTCTGGCAACAGGCTCCCTGACCATCTTGGGCTCCAGGC 300  
 301 TGAGGAGAGGCTGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 357  
 301 TGAGGAGAGGCTGATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 358 CGGAAGTGGACCAAGGTCAAGTCTGAGTCAAGCCCAAGGCAAGCCCACTGCTCACT 417  
 361 CGGAAGTGGACCAAGGTCAAGTCTGAGTCAAGCCCAAGGCAAGCCCACTGCTCACT 420  
 418 GTTCCGCGCTCTCTTGAAGAGTCCCAAGCAACAGGCAACAGTGTGTGTATTCAG 477  
 421 GTTCCGCGCTCTCTTGAAGAGTCCCAAGCAACAGGCAACAGTGTGTGTATTCAG 480

478 TGACTTCAACCCGGAGCTGTGACAGTGGCTGGAAGGCAAGTGGCAAGCCCTGTCAGGC 537  
 481 TGACTTCAACCCGGAGCTGTGACAGTGGCTGGAAGGCAAGTGGCAAGCCCTGTCAGGC 540  
 538 GGAAGTGGAGCAACCAACCTCCCAACAGAGCAACCAAGTGGCAAGCCCTGTCAGGC 597  
 541 GGAAGTGGAGCAACCAACCTCCCAACAGAGCAACCAAGTGGCAAGCCCTGTCAGGC 596  
 598 CTTGAGCTTACAGCCGCTGAGAGTGGAGTCCCAAGAGTCAAGTCTGCTGCTGCTGCTG 657  
 597 CTTGAGCTTACAGCCGCTGAGAGTGGAGTCCCAAGAGTCAAGTCTGCTGCTGCTGCTG 655  
 658 TGAAGGAGCAACCGTGTGATGAGACAGTGTGTC 690  
 658 TGAAGGAGCAACCGTGTGATGAGACAGTGTGTC 688

RESULT 8  
 BM007694 808 bp mRNA linear EST 30-OCT-2001  
 LOCUS 60361712671 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:5440855 5',  
 DEFINITION mRNA sequence.  
 ACCESION BM007694  
 VERSION BM007694.1 GI:16522048  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 808)  
 NIH-MGC http://mgs.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Dr. Mark Watson  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LNCM1914 row: d column: 08  
 High quality sequence stop: 805.  
 Location/Qualifiers  
 1..808  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5440855"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_id="NIH MGC\_113"  
 /note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAAGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

ORIGIN  
 Query Match 41.9%; Score 612 4; DB 12; Length 808;  
 Best Local Similarity 93.9%; Pred. No. 2.2e-129;  
 Matches 648; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

1 GGCCTGGGCTCTGTATTCCTCAACCTCCTCACTCAAGGACAGAGGCTCTGGGCCCAATC 60  
 33 GGCTGGGCTCTGTATTCCTCAACCTCCTCACTCAAGGACAGAGGCTCTGGGCCCAATC 92  
 61 TGCCCTGACTCAAGTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

Db 93 TGCCCTGACTACAGCTCGCTCAGTGTCCGGGTCTCTGGAGACATGACCATCTCTG 152  
Qy 121 CACTGGAACCAAGACAGTGTGTGTATTAATCTATCTCTCTGTATCAACAGACACC 180  
Db 153 CACTGGAACCAAGACAGTGTGTGTATTAATCAACTGTCTCTGTATCAACAGACACC 212  
Qy 181 AGGCAAGGCCCCCAATCATCTATTATAGTCACTAATCGGCTCAGGGGTTCTTA 240  
Db 213 AGGCAAGGCCCCCAATCATCTATTATAGTCACTAATCGGCTCAGGGGTTCTTA 272  
Qy 241 TCGCTTCTCTGTGTCCAGTCTGGCAACAGGCTCTCTGACATCTCTGGCTCAGGC 300  
Db 273 TCGCTTCTCTGTGTCCAGTCTGGCAACAGGCTCTCTGACATCTCTGGCTCAGGC 332  
Qy 301 TGAGAGAGAGGCTGATTATCTGCTGCTCATATACAGAGTACTTCTCATGTCTGG 360  
Db 333 TGAGAGAGAGGCTGATTATCTGCTGCTCATATAGAGCACTACCTATGTCTGG 392  
Qy 361 AACTGGAGACCAAGTCACTGCTCTAGTCAAGCCCAAGCCCAACCCCACTGCTCTGT 420  
Db 393 AACTGGAGACCAAGTCACTGCTCTAGTCAAGCCCAAGCCCAACCCCACTGCTCTGT 452  
Qy 421 CCGGCTCTCTGAGAGAGTCCAGCAAGCCCAAGCCCAAGTGTCTGATCACTGA 480  
Db 453 CCGGCTCTCTGAGAGAGTCCAGCAAGCCCAAGCCCAAGTGTCTGATCACTGA 512  
Qy 481 CTCTTACCCGGAGAGTGTGACAGTGTCTGAGAGAGAGTGTGAGAGAGAGAGAG 540  
Db 513 CTCTTACCCGGAGAGTGTGACAGTGTCTGAGAGAGAGTGTGAGAGAGAGAGAG 572  
Qy 541 AGTGGAGACCAACCAACCTCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGT 600  
Db 573 AGTGGAGACCAACCAACCTCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGT 632  
Qy 601 GAGCTGAGAGCCCGGAGAGAGTCCAGAGAGTCAAGTCAAGTCAAGTCAAGTCA 660  
Db 633 GAGCTGAGAGCCCGGAGAGAGTCCAGAGAGTCAAGTCAAGTCAAGTCAAGTCA 692  
Qy 661 AGGAGACACCGTGTGATGAAGACAGTGTCC 690  
Db 693 AGGAGACACCGTGTGATGAAGACAGTGTCC 721

RESULT 9  
LOCUS BM920304 1049 bp mRNA linear EST 12-MAR-2002  
DEFINITION AGENCOURT 5706549 NIH\_MGC\_120 Homo sapiens CDNA clone IMAGE:5750228  
5', mRNA sequence.  
ACCESSION BM920304  
VERSION BM920304.1 GI:19370683  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 1049)  
NIH-MGC <http://mgi.nci.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
[found through the I.M.A.G.E. Consortium/LLNL at:](http://image.llnl.gov)  
<http://image.llnl.gov>  
Plate: LRAM2780 row: n column: 21  
High quality sequence stop: 669.  
Location/Qualifiers  
1..1049  
/organism="Homo sapiens"

ORIGIN  
Query Match 41.8%; Score 610.8; DB 12; Length 1049;  
Best Local Similarity 93.8%; Pred. No. 5,9e-129;  
Matches 636; Conservative 0; Mismatches 42; Indels 0; Gaps 0;  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5750228"  
/lab\_host="DH10B"  
/clone\_1db="NIH MGC 120"  
/note="Organ: pooled pancreas and spleen; Vector:  
pCMV-Sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
source anonymous Pool of spleen and pancreas from 28 yo  
male. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.5 kb, insert size range 1-2.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 025. Note: this is a NIH\_MGC Library."

Qy 1 GAGCTGGAGTCTGCTATTCTCCACCCCTCCATCAAGGACAGAGGCTCTGGGCCAGTC 60  
Db 166 GAGCTGGAGTCTGCTATTCTCTCTCTCACTCTCTCACTCAAGGACAGAGGCTCTGGGCCAGTC 225  
Qy 61 TGCCCTGACTGACTGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
Db 226 TGCTCTGACTGACTGCTCCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285  
Qy 121 CACTGGAACCAAGAGTACAGTGTGTGTATTAATCTATCTCTGCTGCTGCTGCTGCTG 180  
Db 286 CGCTGGAACCAAGAGTACAGTGTGTGTATTAATCAAGTCTCTGCTGCTGCTGCTGCTG 345  
Qy 181 AGGCAAGGCCCCCAATCATCTATTATAGAGTCAAGTCAAGTCAAGTCAAGTCAAG 240  
Db 346 AGGCAAGGCCCCCAATCATCTATTATAGAGTCAAGTCAAGTCAAGTCAAGTCAAG 405  
Qy 241 TCGCTTCTCTGTGTCCAGTGTGCAAGAGGCTCTCTGACATCTCTGGCTCAGGC 300  
Db 406 TCGCTTCTCTGTGTCCAGTGTGCAAGAGGCTCTCTGACATCTCTGGCTCAGGC 465  
Qy 301 TGAGAGAGAGGCTGATTATCTGCTGCTCATATTAAGAGAGTCAAGTCAAGTCAAGTCA 360  
Db 466 TGAGAGAGAGGCTGATTATCTGCTGCTCATATTAAGAGAGTCAAGTCAAGTCAAGTCA 525  
Qy 481 CTCTTACCCGGAGAGTGTGACAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGT 540  
Db 526 AACTGGAGACCAAGTCACTGCTCTAGGTCAAGCCCAAGCCCAACCCCACTGCTCTGT 585  
Qy 421 CCGGCTCTCTGAGAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGTGTGTGATCACTGA 480  
Db 586 CCGGCTCTCTGAGAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGTGTGTGATCACTGA 645  
Qy 481 CTCTTACCCGGAGAGTGTGACAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGT 540  
Db 646 CTCTTACCCGGAGAGTGTGACAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGT 705  
Qy 541 AGTGGAGACCAACCAACCTCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGT 600  
Db 706 AGTGGAGACCAACCAACCTCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGT 765  
Qy 601 GAGCTGAGCCCGGAGAGAGTCCAGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 660  
Db 766 GAGCTGAGCCCGGAGAGAGTCCAGAGAGTCCAGAGAGTCAAGTCAAGTCAAGTCA 825  
Qy 661 AGGAGACACCGTGTGATGA 678  
Db 826 AGGAGACACCGTGTGATGA 843

RESULT 10  
LOCUS BG759257 686 bp mRNA linear EST 15-MAY-2001  
DEFINITION BG759257 602710936F1 NIH\_MGC\_48 Homo sapiens CDNA clone IMAGE:4851128 5',



Db 157 TGCCCTGACTGACGCTGCTCCGCTGCTGGTCTCTCCCTGAGCAGTCGATCACCATCTCTG 216

Qy 121 CACTGGAACAGACAGACAGCTGGTGGTTATTAATCTGTCTCTCTGTACCAACAGACACC 180

Db 217 CACTGGAACAGACAGACAGCTGGTGGTTATTAATCTGTCTCTGTACCAACAAACC 276

Qy 181 AGGCAAGAGCCCAAACTCATCATTTATGAGTCAAGTATCGAGCCCTCAGGGTTCTTA 240

Db 277 AGGCAAGAGCCCAAACTCATCATTTATGAGTCAAGTATCGAGCCCTCAGGGATTCTTA 336

Qy 241 TGCTTCTCTGCTGCTCAAGTCTGGCAACAGGCTCTCTGACCACTCTGAGCTCAAGC 300

Db 337 TGCTTCTCTGCTGCTCAAGTCTGGCAACAGGCTCTCTGACCACTCTGAGCTCAAGC 396

Qy 301 TGAGACAGAGCTGATTATTAATCTGCTGCTCATATACAA---GAAGTACTCTCATGTCT 357

Db 397 TGAGACAGAGCTGATTATTAATCTGCTGCTCATATACAAACAGCAATCTCTCAAGTAT 456

Qy 358 CGGAAGTGGAGACCAAGCTACCGCTCTAGAGTCAAGCCCAAGCCCACTGCTACTCT 417

Db 457 CGGCGAGGAGACCAAGCTACCGCTCTAGAGTCAAGCCCAAGCCCTGCTGCTACTCT 516

Qy 418 GTTCCGCGCCCTCTCTGAGAGTCCCAAGCCCAAGCCCAAGTGTCTGATCAG 477

Db 517 GTTCCGCGCCCTCTCTGAGAGTCCCAAGCCCAAGCCCAAGTGTCTGATCAG 576

Qy 478 TGAATCTTACCCGAGAGCTGTGACAGTGGCTGGAAGCAGATGGCAGCCCGTCAAGC 537

Db 577 TGAATCTTACCCGAGAGCTGTGACAGTGGCTGGAAGCAGATGGCAGCCCGTCAAGC 636

Qy 538 GGGAGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAAGTACGGGCGCAGAGCTA 597

Db 637 GGGAGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAAGTACGGGCGCAGAGCTA 696

Qy 598 CTGAGACCTGACGCGCCGAGAGTGAAGTCCCAAGAGCTACAGTGGCAGAGTCAAGCA 657

Db 697 CTGAGACCTGACGCGCCGAGAGTGAAGTCCCAAGAGCTACAGTGGCAGAGTCAAGCA 756

Qy 658 TGAAGGAGACCAAGTGTATGAAGACAGTGTCC 691

Db 757 TGAAGGAGACCAAGTGTATGAAGACAGTGTCC 789

RESULT 12 1150 bp mRNA linear EST 22-MAY-2003

LOCUS BX449367

DEFINITION BX449367 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA

ACCESSION BX449367

VERSION BX449367.1 GI:31032034

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Li, M.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: sequef@genoscope.cns.fr Web: www.genoscope.cns.fr

Library was constructed by life technologies, a division of invitrogen. This sequence belongs to sequence cluster 10531.f For more information about this cluster, see

http://www.genoscope.cns.fr/

cgf-bin/cluster.cgi?seq=CS0AP003ZB11QPI&cluster=10531.f. Contact :

Feng Liang Email : fliang@life.techn.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0AP003ZB11QPI.

Location/Qualifiers

1. 1150

FEATURES

source

Query Match 41.6%; Score 606.8; DB 13; Length 1150;

Best Local Similarity 94.1%; Pred. No. 51e-128;

Matches 653; Conservative 0; Mismatches 37; Indels 4; Gaps 2;

ORIGIN

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D002YH06"

/issue\_type="B CELLS (RAMOS CELL LINE)"

/cell\_line="RAMOS CELL LINE"

/clone\_id="Homo sapiens B CELLS (RAMOS CELL LINE)"

/note="Vector: PCWVSPT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the PCWVSPT 6 vector.

Library was not normalized."

1 GGCCTGGGCTCTGCTATTCCTCAACCTCTCACTCAGGAGCAGAGTCTGGGCCAGTC 60

97 GGCCTGGGCTCTGCTATTCCTCAACCTCTCACTCAGGAGCAGAGTCTGGGCCAGTC 156

Qy 61 TGCCCTGACTGAGTTCGCTCTCGTGTGGGTCTCTCTGAGCAGTCCATCACTTCCTG 120

Db 157 TGCCCTGACTGAGTTCGCTCTCGTGTGGGTCTCTCTGAGCAGTCCATCACTTCCTG 216

Qy 121 CACTGGAACAGACAGTACGCTGGTGGTTATTAATGATGCTCTGTTACCAAGACACC 180

Db 217 CACTGGAACAGACAGTACGCTGGTGGTTATTAATGATGCTCTGTTACCAAGACACC 276

Qy 181 AGGCAAGAGCCCAAACTCATCATTTATGAGTCAAGTATCGAGCCCTCAGGGTTCTTA 240

Db 277 AGGCAAGAGCCCAAACTCATCATTTATGAGTCAAGTATCGAGCCCTCAGGGTTCTTA 336

Qy 241 TGCTTCTCTGCTGCTCAAGTCTGGCAACAGGCTCTCTGACCACTCTGAGCTCAAGC 300

Db 337 TGCTTCTCTGCTGCTCAAGTCTGGCAACAGGCTCTCTGACCACTCTGAGCTCAAGC 396

Qy 301 TGAGACAGAGCTGATTATTAATCTGCTGCTCATATACAA---GAAGTACTCTCATGTCT 357

Db 397 TGAGACAGAGCTGATTATTAATCTGCTGCTCATATACAAACAGCAATCTCTCAAGTAT 456

Qy 358 CGGAAGTGGAGACCAAGTACCGCTCTAGAGTCAAGCCCAAGCCCACTGCTACTCT 417

Db 457 CGGCGAGGAGACCAAGTACCGCTCTAGAGTCAAGCCCAAGCCCACTGCTACTCT 516

Qy 418 GTTCCGCGCCCTCTCTGAGAGTCCCAAGCCCAAGCCCAAGTGTCTGATCAG 477

Db 517 GTTCCGCGCCCTCTCTGAGAGTCCCAAGCCCAAGCCCAAGTGTCTGATCAG 576

Qy 478 TGAATCTTACCCGAGAGCTGTGACAGTGGCTGGAAGCAGATGGCAGCCCGTCAAGC 537

Db 577 TGAATCTTACCCGAGAGCTGTGACAGTGGCTGGAAGCAGATGGCAGCCCGTCAAGC 636

Qy 538 GGGAGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAAGTACGGGCGCAGAGCTA 597

Db 637 GGGAGTGGAGACCAACCAACCTCTCCAAACAGAGCAACAAAGTACGGGCGCAGAGCTA 696

Qy 598 CTGAGACCTGACGCGCCGAGAGTGAAGTCCCAAGAGCTACAGTGGCAGAGTCAAGCA 657

Db 697 CTGAGACCTGACGCGCCGAGAGTGAAGTCCCAAGAGCTACAGTGGCAGAGTCAAGCA 756

Qy 658 TGAAGGAGACCAAGTGTATGAAGACAGTGTCC 691

Db 757 TGAAGGAGACCAAGTGTATGAAGACAGTGTCC 789

RESULT 13 1201 bp mRNA linear EST 22-MAY-2003

LOCUS BX462166

DEFINITION BX462166 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA

ACCESSION BX462166

clone CS0D002YH06 5-PRIME, mRNA sequence.

VERSION BX462166.1 GI:3103379  
 EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 1201)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10531.f. For more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DG002D03QPI&cluster=10531.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Paraday Avenue Genoscope sequence ID : CS0DG002D03QPI.  
 Location/Qualifiers  
 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DG002YH06"  
 /rname="B CELLS (RAMOS CELL LINE)"  
 /cell\_line="RAMOS CELL LINE"  
 /clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"  
 /note="Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the PCMVSPORT 6 vector. Library was not normalized."

FEATURES  
 source  
 1..1201

## ORIGIN

Query Match 41.5%; Score 606.4; DB 13; Length 1201;  
 Best Local Similarity 93.9%; Pred. No. 6,4e-128;  
 Matches 652; Conservative 1; Mismatches 37; Indels 4; Gaps 2;

```

1  GGGCTGGGCTGCTGCTATTCCTCACTCCTCACTCAAGGACAGAGGCTCTGGAGCCCAAGC 60
86  GGGCTGGGCTGCTGCTATTCCTCACTCCTCACTCAAGGACAGAGGCTCTGGAGCCCAAGC 145
61  TGGCTGACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
146  TGGCTGACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 205
121  CAGTGAACAGCAGCAGTCAAGTGTGTTATTAATGTTCTCTGCTGCTGCTGCTGCTGCTGCT 180
206  CAGTGAACAGCAGCAGTCAAGTGTGTTATTAATGTTCTCTGCTGCTGCTGCTGCTGCTGCT 265
181  AGGCAAGCCCAAACTCATTTATGAGTCACTAAATGAGCCCTCAGAGGATTTCTAA 240
266  AGGCAAGCCCAAACTCATTTATGAGTCACTAAATGAGCCCTCAGAGGATTTCTAA 325
241  TGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
326  TGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 385
301  TGAAGAGAGGCTGATTTATGCTGCTCATATACAA--GAAGTACTTCTCATGCTCT 357
386  TGAAGAGAGGCTGATTTATGCTGCTCATATATACAAAGAGAGATTTCTCATGCTCT 445
446  CGAGAGAGGAGCAAGTGAAGCTCTAGTCAAGCCCAAGGCTGCTGCTGCTGCTGCTGCT 505
418  GTTCCGAGCTCTCTGAGAGGCTCAAGGCAACAGAGGCAACAGGCTGCTGCTGCTGCTGCT 477
506  GTTCCGAGCTCTCTGAGAGGCTCAAGGCAACAGAGGCAACAGGCTGCTGCTGCTGCTGCT 565

```

```

478  TGAAGAGAGGCTGATTTATGCTGCTCATATACAA--GAAGTACTTCTCATGCTCT 537
566  TGAAGAGAGGCTGATTTATGCTGCTCATATATACAAAGAGAGATTTCTCATGCTCT 625
538  GGAAGTGAAGCAACCAACCTCTCAACAGAGCAACAGTACAGGCTGCTGCTGCTGCTGCT 597
626  GGAAGTGAAGCAACCAACCTCTCAACAGAGCAACAGTACAGGCTGCTGCTGCTGCTGCT 685
598  CCTGAGCTGAGCGCCGAGAGAGTGAAGTCCCAAGAGCTACAGTCCAGGCTGCTGCTGCT 657
686  CCTGAGCTGAGCGCCGAGAGAGTGAAGTCCCAAGAGCTACAGTCCAGGCTGCTGCTGCT 745
746  TGAAGAGAGCAAGCTGGA--GAAGCAAGTGGCCCC 778

```

## RESULT 14

EX442894

LOCUS

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10531.f. For more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DG001AF09QPI&cluster=10531.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Paraday Avenue Genoscope sequence ID : CS0DG001AF09QPI.  
 Location/Qualifiers  
 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DG001YK17"  
 /rname="B CELLS (RAMOS CELL LINE)"  
 /cell\_line="RAMOS CELL LINE"  
 /clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE)"  
 /note="Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the PCMVSPORT 6 vector. Library was not normalized."

## FEATURES

source

## ORIGIN

Query Match 41.4%; Score 604; DB 13; Length 1201;  
 Best Local Similarity 93.5%; Pred. No. 2,3e-127;  
 Matches 649; Conservative 2; Mismatches 39; Indels 4; Gaps 2;

```

1  GGGCTGGGCTGCTGCTATTCCTCACTCCTCACTCAAGGACAGAGGCTCTGGAGCCCAAGC 60
100  GGGCTGGGCTGCTGCTATTCCTCACTCCTCACTCAAGGACAGAGGCTCTGGAGCCCAAGC 159
61  TGGCTGACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
160  TGGCTGACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 219

```



